79576

Fr m: Chan, Christina
Sent: Wednesday, November 06, 2002 1:16 PM
To: Holleran, Anne, STIC-Biotech/ChemLib
Subject: RE: RUSH search for 09/480,977

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

Point of Contact:
Mona Smith
Technical Information Specialist
CM1 6A01
Tel: 308-3278

-----Original Message-----
From: Holleran, Anne
Sent: Wednesday, November 06, 2002 9:52 AM
To: Chan, Christina
Subject: RUSH search for 09/480,977

Please approve and forward to STIC the following sequence search request. This case is an amendment due this biweek. Thanks.

Please search the following:

commercial and interference database search of SEQ ID NO: 4(aa)

Anne Holleran
AU: 1642
Tel: 308-8892
RM: 8e03

mailbox: 8e12

Searcher: M. Smith
Phone:_____
Location:_____
Date Picked Up: 11 7 02
Date Completed: 11 8 02
Searcher Prep/Review:___5
Clerical:_____
Online time:_____5

TYPE OF SEARCH:
NA Sequences:_____
AA Sequences:_____1
Structures:_____
Bibliographic:_____
Litigation:_____
Full text:_____
Patent Family:_____
Other:_____

VENDOR/COST (where applic.)
STN:_____
DIALOG:_____
Questel/Orbit:_____
DRLink:_____
Lexis/Nexis:_____
Sequence Sys.:_____
WWW/Internet:_____
Other (specify):_____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 7, 2002, 09:59:10 ; Search time 26 seconds
(without alignments)
312.722 Million cell updates/sec

Title:
Perfect score: 277
Sequence: 1 HPKPCRDKDLAVCLNDGECF.........SHKHCRCKECYQGVRCDQFL 47   US-09-480-977-4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_virus:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	113.5	41.0	461	11	O35947	O35947 mesocricetu
2	111.5	40.3	298	11	Q9esa9	Q9esa9 rattus norv
3	111.5	40.3	695	11	Q9esb0	Q9esb0 rattus norv
4	104.5	37.7	241	6	O07112	O07112 bos taurus
5	104.5	37.7	296	4	Q96IB3	Q96ib3 homo sapien
6	102.5	37.0	111	11	Q9ESA8	Q9esa8 rattus norv
7	102.5	37.0	136	11	Q9ESA7	Q9esa7 rattus norv
8	102.5	37.0	256	11	Q9ESA6	Q9esa6 rattus norv
9	102.5	37.0	317	11	Q9ESA5	Q9esa5 rattus norv
10	102.5	37.0	323	11	Q9ESA3	Q9esa3 rattus norv
11	102.5	37.0	342	11	Q9ESA2	Q9esa2 rattus norv
12	102.5	37.0	700	11	Q9ESA1	Q9esa1 rattus norv
13	102.5	37.0	782	11	Q9ESB1	Q9esb1 rattus norv
14	99	35.7	89	12	Q91MZ0	Q91mz0 lumpy skin
15	92.5	33.4	2192	5	O01768	O01768 caenorhabdi
16	91.5	33.0	1213	13	Q90Y54	Q90y54 brachydanio

ALIGNMENTS

RESULT 1
O35947
ID   O35947   PRELIMINARY;   PRT;   461 AA.
AC   O35947;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PRO-NEUREGULIN-1, ISOFORM ALPHA 2B PRECURSOR.
GN   NRG1 OR NDF.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA2B), AND SEQUENCE OF 64-81.
RC   TISSUE=EMBRYO;
RX   MEDLINE=98196996; PubMed=9537646;
RA   Velasco J.A., Feijoo E., Avila M.A., Notario V.;
RT   "Secretion of neu differentiation factor-like polypeptides by cph-
RT   transformed fibroblasts: cloning and characterization of Syrian
RT   hamster neuregulin cDNAs.";
RL   Mol. Carcinog. 21:156-163(1998).
CC   -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC   RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC   RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC   ACTIVATION OF THE ERBB RECEPTORS. MAY PLAY AN IMPORTANT ROLE IN
CC   PROVIDING GROWTH ADVANTAGE IN NEOPLASTIC CELLS.
CC   -!- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
CC   REGION OF LIMK1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: EXISTS AS TYPE I MEMBRANE PROTEIN AND AS A
CC   PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
CC   BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS ARE PRODUCED BY
CC   ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM
CC   ALPHA2B/CLONE PM3.
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHER LEVEL AFTER NEOPLASMIC
CC   TRANSFORMATION OF CELLS.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION

17	89	32.1	162	11	Q9z0L5
18	89	32.1	1241	4	Q9UKK5
19	89	32.1	1241	4	Q9UKD4
20	89	32.1	1241	4	Q9BZV3
21	89	31.8	1239	11	P70628
22	87	31.4	217	5	Q9VVJ6
23	87	31.4	597	11	O35727
24	85.5	30.9	1193	13	Q90819
25	85.5	30.9	1214	13	Q90D2
26	85.5	30.9	1218	4	O14902
27	85.5	30.9	1218	4	O15122
28	85.5	30.9	1218	4	O15816
29	85.5	30.9	1227	4	P78504
30	84.5	30.5	177	11	Q9JJM4
31	84.5	30.5	1218	11	Q9QXX0
32	84.5	30.5	1219	11	O63722
33	83.5	30.1	127	12	Q9Q9F3
34	83	30.0	162	11	Q61521
35	83	30.0	861	11	Q9QW58
36	83	30.0	862	11	Q99JC2
37	82.5	29.8	445	5	Q9W3W5
38	82.5	29.8	456	11	Q961F3
39	82	29.6	264	5	Q63722
40	81	29.2	78	12	Q91T36
41	80.5	29.1	125	12	O41504
42	80.5	29.1	178	4	Q96F48
43	80.5	29.1	260	4	Q95898
44	80.5	29.1	530	4	Q96CW9
45	80.5	29.1	1242	13	Q90Y57

Q9z0L5 rattus norv
Q9UKK5 homo sapien
Q9UKD4 homo sapien
Q9BZV3 homo sapien
P70628 rattus norv
Q9VVJ6 drosophila
O35727 mus musculu
Q90819 gallus gall
Q90D2 xenopus lae
O14902 homo sapien
O15122 homo sapien
O15816 homo sapien
P78504 homo sapien
Q9JJM4 rattus norv
Q9QXX0 mus musculu
O63722 rattus norv
Q9Q9F3 ectromelia
Q61521 mus musculu
Q9QW58 mus sp. mot
Q99JC2 mus musculu
Q9W3W5 drosophila
Q961F3 drosophila
O63722 caenorhabdi
Q91T36 lumpy skin
O41504 cowpox viru
Q96F48 homo sapien
Q95898 homo sapien
Q96CW9 homo sapien
Q90Y57 brachydanio

```
CC        OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC        PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC        DIMERIZATION (BY SIMILARITY).
CC   -!-  DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC        DOMAIN.
CC   -!-  PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC        EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC        FORM (BY SIMILARITY).
CC   -!-  PTM: EXTENSIVE GLYCOSYLATION PRECEDES PROTEOLYTIC CLEAVAGE (BY
CC        SIMILARITY).
CC   -!-  SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!-  SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!-  SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
DR        EMBL; Q12784; 1HRE.
DR        HSSP; Q12784; 1HRE.
DR        InterPro; IPR000561; EGF-like.
DR        InterPro; IPR003598; Ig_c2.
DR        InterPro; IPR003006; Ig_MHC.
DR        InterPro; IPR002154; Neuregulin.
DR        Pfam; PF00008; EGF; 1.
DR        Pfam; PF00047; Ig; 1.
DR        Pfam; PF02158; Neuregulin; 1.
DR        PRINTS; PR01089; NEUREGULIN.
DR        SMART; SM00181; EGF; 1.
DR        SMART; SM00408; IGc2; 1.
DR        PROSITE; PS00022; EGF_1; 1.
DR        PROSITE; PS01186; EGF_2; 1.
KW        Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW        Transmembrane; Alternative splicing.
FT        PROPEP          1      13       PRO-NEUREGULIN-1.
FT        CHAIN          14     461       NEUREGULIN-1, MEMBRANE-BOUND FORM.
FT        CHAIN          14     241       NEUREGULIN-1.
FT        DOMAIN         14     242       EXTRACELLULAR (POTENTIAL).
FT        TRANSMEM      243     265       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT        DOMAIN        266     461       CYTOPLASMIC (POTENTIAL).
FT        DOMAIN         50     119       IG-LIKE C2-TYPE DOMAIN.
FT        DOMAIN        165     177       SER/THR-RICH.
FT        DOMAIN        178     222       EGF-LIKE.
FT        DISULFID       57     112       BY SIMILARITY.
FT        DISULFID      182     196       BY SIMILARITY.
FT        DISULFID      190     210       BY SIMILARITY.
FT        DISULFID      212     221       BY SIMILARITY.
FT        CARBOHYD       73      73       N-LINKED (GLCNAC...) (POTENTIAL).
FT        CARBOHYD      120     120       N-LINKED (GLCNAC...) (POTENTIAL).
FT        CARBOHYD      126     126       N-LINKED (GLCNAC...) (POTENTIAL).
FT        CARBOHYD      164     164       N-LINKED (GLCNAC...) (POTENTIAL).
SQ        SEQUENCE   461 AA;  50890 MW;  935C95607F148336 CRC64;
```

```
Query Match              41.0%;  Score 113.5;  DB 11;  Length 461;
Best Local Similarity    34.8%;  Pred. No. 1e-07;
Matches   16;  Conservative   14;  Mismatches  15;  Indels   1;  Gaps   1;

Qy        1 HFKPCRDKDLAYCLNDGECFVIETLIGSHKH-CRCKEGYQGVRCDQ 45

Db      178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 223
```

RESULT 2

```
Q9ESA9
ID   Q9ESA9          PRELIMINARY;      PRT;   298 AA.
AC   Q9ESA9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMDF NEUREGULIN ALPHA 2B (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDIX;
RA   Carroll S.L., Anderson K.D., Frohnert P.W.;
RT   "Structural and Functional Diversity of SMDF Neuregulin Splice
RT   Variants Expressed in the Adult Rat Nervous System.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194440; AAG28429.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002154; Neuregulin.
DR   InterPro; IPR002114; PTS_HPr_ser.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
FT   NON_TER        1       1
FT   NON_TER      298     298
SQ   SEQUENCE  298 AA;  32851 MW;  BD76F014C2B33026 CRC64;
```

```
Query Match              40.3%;  Score 111.5;  DB 11;  Length 298;
Best Local Similarity    34.8%;  Pred. No. 1.3e-07;
Matches   16;  Conservative   13;  Mismatches  16;  Indels   1;  Gaps   1;

Qy        1 HFKPCRDKDLAYCLNDGECFVIETLIGSHKH-CRCKEGYQGVRCDQ 45

Db       48 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCQPGFTGARCTE 93
```

RESULT 3

```
Q9ESB0
ID   Q9ESB0          PRELIMINARY;      PRT;   695 AA.
AC   Q9ESB0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMDF NEUREGULIN ALPHA 2A.
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDIX;
RA   Carroll S.L., Anderson K.D., Frohnert P.W.;
RT   "Structural and Functional Diversity of SMDF Neuregulin Splice
RT   Variants Expressed in the Adult Rat Nervous System.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194439; AAG28428.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002154; Neuregulin.
DR   InterPro; IPR002114; PTS_HPr_ser.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
SQ   SEQUENCE  695 AA;  75646 MW;  5277F2CBA2FB6878 CRC64;
```

```
Query Match              40.3%;  Score 111.5;  DB 11;  Length 695;
Best Local Similarity    34.8%;  Pred. No. 3.1e-07;
Matches   16;  Conservative   13;  Mismatches  16;  Indels   1;  Gaps   1;

Qy        1 HFKPCRDKDLAYCLNDGECFVIETLIGSHKH-CRCKEGYQGVRCDQ 45

Db      234 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCQPGFTGARCTE 279
```

RESULT 4

```
ID   Q07112      PRELIMINARY;   PRT;   241 AA.
AC   Q07112;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   NEUREGULIN-1, GLIAL GROWTH FACTOR 5 ISOFORM PRECURSOR (GGFBPP5).
GN   NRG1 OR GGF.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC   TISSUE=POSTERIOR PITUITARY;
RX   MEDLINE=93205115; PubMed=8096067;
RA   Marchionni M.A., Goodearl A.D.J., Chen M.S., Bermingham-McDonogh O.,
RA   Kirk C., Hendricks M., Danehy F., Misumi D., Sudhalter J.,
RA   Kobayashi K., Wroblewski D., Lynch C., Baldassare M., Hiles I.,
RA   Davis J.B., Hsuan J.J., Totty N.F., Otsu M., McBurney R.N.,
RA   Waterfield M.D., Stroobant P., Gwynne D.;
RT   "Glial growth factors are alternatively spliced erbB2 ligands
RT   expressed in the nervous system.";
RL   Nature 362:312-318(1993).
CC   -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC       RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC       RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE ERBB RECEPTORS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD AND
CC       BRAIN.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST FIVE DIFFERENT ISOFORMS; GGFBPP1,
CC       GGFBPP2, GGFBPP3, GGFBPP4 AND GGFBPP5 (SHOWN HERE); ARE PRODUCED
CC       BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
DR   EMBL; L12259; AAA30540.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   Growth factor; EGF-like domain; Immunoglobulin domain;
KW   Alternative splicing.
FT   PROPEP        1     19       BY SIMILARITY.
FT   CHAIN        20    241       NEUREGULIN-1, GLIAL GROWTH FACTOR 5
FT                                ISOFORM.
FT   DOMAIN       50    119       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      165    177       SER/THR-RICH.
FT   DOMAIN      178    222       EGF-LIKE DOMAIN.
FT   DISULFID    182    196       BY SIMILARITY.
FT   DISULFID    190    210       BY SIMILARITY.
FT   DISULFID    212    221       BY SIMILARITY.
FT   DISULFID    134    156       BY SIMILARITY.
FT   VARSPLIC    157    241       EITTGMPASTETAYVSSESPIRI -> GKRCLLRAISQSLR
FT                                GVIKVCGHT (IN ISOFORM GGFBPP1).
FT   VARSPLIC    157    156       MISSING (IN ISOFORM GGFBPP1).
SQ   SEQUENCE   241 AA;  25955 MW;  BF571297E8DA9796 CRC64;

Qy  Query Match
    Best Local Similarity   37.7%;    Score 104.5; DB 6;  Length 241;
    Matches 15;  Conservative 14;  Mismatches 18;  Indels 1;  Gaps 1;
                                     Pred. No. 9.7e-07;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         | ||: :|:| |||||| |:|:|| |::|| :|:: | | | :
Db  178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCPNEFTGDRCQNYV 225
```

```
RESULT 5
ID   Q96IB3      PRELIMINARY;   PRT;   296 AA.
AC   Q96IB3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NEUREGULIN 1.
GN   NRG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND NEUROBLASTOMA;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC007675; AAH07675.1; -.
SQ   SEQUENCE   296 AA;  31699 MW;  0F5C48C84656B649 CRC64;

Qy  Query Match
    Best Local Similarity   37.7%;    Score 104.5; DB 4;  Length 296;
    Matches 15;  Conservative 14;  Mismatches 18;  Indels 1;  Gaps 1;
                                     Pred. No. 1.2e-06;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         | ||: :|:| |||||| |:|:|| |::|| :|:: | | | :
Db  233 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCPNEFTGDRCQNYV 280
```

```
RESULT 6
ID   Q9ESA8      PRELIMINARY;   PRT;   111 AA.
AC   Q9ESA8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMDF NEUREGULIN BETA 2 (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RC   TISSUE=AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA   Carroll S.L., Anderson K.D., Frohnert P.W.;
RT   "Structural and Functional Diversity of SMDF Neuregulin Splice
RT   Variants Expressed in the Adult Rat Nervous System.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ database.
DR   EMBL; AF194441; AAG28430.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR0008886; ER_target.
DR   Pfam; PF00008; EGF; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     111    111
SQ   SEQUENCE   111 AA;  12198 MW;  CC8B88705B4C9F8C CRC64;

Qy  Query Match
    Best Local Similarity   37.0%;    Score 102.5; DB 11;  Length 111;
    Matches 15;  Conservative 13;  Mismatches 19;  Indels 1;  Gaps 1;
                                     Pred. No. 8.1e-07;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         | ||: :|:| |||||| |:|:|| |::|| :|:: | | | :
Db   54 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 101
```

RESULT 7

```
ID   Q9ESA7      PRELIMINARY;   PRT;   136 AA.
AC   Q9ESA7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMDF NEUREGULIN BETA 4 (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Carroll S.L., Anderson K.D., Frohnert P.W.;
RT   "Structural and Functional Diversity of SMDF Neuregulin Splice
RT   Variants Expressed in the Adult Rat Nervous System.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194442; AAG28431.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   Pfam; PF00008; EGF; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     136    136
SQ   SEQUENCE   136 AA;  15235 MW;  116CBC91D6E8AEF5 CRC64;

Query Match             37.0%;
Best Local Similarity   31.2%;
Matches   15;  Conservative  13;  Mismatches  19;  Indels  1;  Gaps  1;
Score 102.5;  DB 11;  Length 136;  Pred. No. 1e-06;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         ||:|: |:| | |||| || |:   :| || | |:|:|| | | ::
Db   54 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 101
```

```
RESULT 8
ID   Q9ESA6      PRELIMINARY;   PRT;   256 AA.
AC   Q9ESA6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMDF NEUREGULIN BETA 3 (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RC   TISSUE=AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA   Carroll S.L., Anderson K.D., Frohnert P.W.;
RT   "Structural and Functional Diversity of SMDF Neuregulin Splice
RT   Variants Expressed in the Adult Rat Nervous System.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194443; AAG28432.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   Pfam; PF00008; EGF; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     256    256
SQ   SEQUENCE   256 AA;  27335 MW;  C8C08DEA68D33E39 CRC64;

Query Match             37.0%;
Best Local Similarity   31.2%;
Matches   15;  Conservative  13;  Mismatches  19;  Indels  1;  Gaps  1;
Score 102.5;  DB 11;  Length 256;  Pred. No. 2e-06;
```

```
Matches   15;  Conservative  13;  Mismatches  19;  Indels  1;  Gaps  1;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         ||:|: |:| | |||| || |:   :| || | |:|:|| | | ::
Db  193 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 240
```

```
RESULT 9
ID   Q9ESA3      PRELIMINARY;   PRT;   317 AA.
AC   Q9ESA3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLIAL GROWTH FACTOR GGF BETA 2 (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RC   TISSUE=AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA   Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,
RA   Frohnert P.W.;
RT   "Structural and Functional Diversity of Glial Growth Factor Isoforms
RT   Expressed in Regenerating Peripheral Nerve and Associated Neurons.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194995; AAG28449.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000886; ER_TARGET.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig-c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW   Immunoglobulin domain.
FT   NON_TER       1      1
FT   NON_TER     317    317
SQ   SEQUENCE   317 AA;  34785 MW;  4487FA3E9CD876B9 CRC64;

Query Match             37.0%;
Best Local Similarity   31.2%;
Matches   15;  Conservative  13;  Mismatches  19;  Indels  1;  Gaps  1;
Score 102.5;  Pred. No. 2.5e-06;  DB 11;  Length 317;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         ||:|: |:| | |||| || |:   :| || | |:|:|| | | ::
Db  260 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 307
```

```
RESULT 10
ID   Q9ESA2      PRELIMINARY;   PRT;   323 AA.
AC   Q9ESA2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLIAL GROWTH FACTOR GGF BETA 3 (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RA   Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,
```

```
RESULT 11
ID   Q9ESA1      PRELIMINARY;      PRT;    342 AA.
AC   Q9ESA1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLIAL GROWTH FACTOR GGF BETA 4 (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RC   TISSUE=AXOTOMIED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA   Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,
RA   Frohnert P.W.;
RT   "Structural and Functional Diversity of Glial Growth Factor Isoforms
RT   Expressed in Regenerating Peripheral Nerve and Associated Neurons.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194997; AAG28451.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW   Immunoglobulin domain.
FT   NON_TER       1
FT   NON_TER     342
SQ   SEQUENCE   342 AA;  37836 MW;  8BE36FC83655J124 CRC64;

Query Match             37.0%;  Score 102.5; DB 11;  Length 342;
Best Local Similarity   31.2%;  Pred. No. 2.7e-06;
Matches    15;  Conservative    13;  Mismatches    19;  Indels     1;  Gaps     1;
```

```
RA   Frohnert P.W.;
RT   "Structural and Functional Diversity of Glial Growth Factor Isoforms
RT   Expressed in Regenerating Peripheral Nerve and Associated Neurons.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194996; AAG28450.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW   Immunoglobulin domain.
FT   NON_TER       1
FT   NON_TER     323
SQ   SEQUENCE   323 AA;  35558 MW;  C7DF153A939A80C8 CRC64;

Query Match             37.0%;  Score 102.5; DB 11;  Length 323;
Best Local Similarity   31.2%;  Pred. No. 2.5e-06;
Matches    15;  Conservative    13;  Mismatches    19;  Indels     1;  Gaps     1;

Qy      1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
           ||||||||||| :|:| ||||| |: |  ||| ||| ||  |||  ::
Db    260 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 307
```

```
RESULT 12
ID   Q9ESB1      PRELIMINARY;      PRT;    700 AA.
AC   Q9ESB1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMDF NEUREGULIN BETA 1A.
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RA   Carroll S.L., Anderson K.D., Frohnert P.W.;
RT   "Structural and Functional Diversity of SMDF Neuregulin Splice
RT   variants Expressed in the Adult Rat Nervous System.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194438; AAG28427.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002154; Neuregulin.
DR   InterPro; IPR002114; PTS_HPr_ser.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS005889; PTS_HPR_SER; UNKNOWN_1.
SQ   SEQUENCE   700 AA;  76386 MW;  2F8111B17ECC49DA CRC64;

Query Match             37.0%;  Score 102.5; DB 11;  Length 700;
Best Local Similarity   31.2%;  Pred. No. 5.7e-06;
Matches    15;  Conservative    13;  Mismatches    19;  Indels     1;  Gaps     1;

Qy      1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
           ||||||||||| :|:| ||||| |: |  ||| ||| ||  |||  ::
Db    234 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 281
```

```
RESULT 13
ID   Q9ESA5      PRELIMINARY;      PRT;    782 AA.
AC   Q9ESA5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLIAL GROWTH FACTOR BETA 1A (FRAGMENT).
GN   NRG1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=SPINAL CORD/BRAIN STEM;
RA   Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,
RA   Frohnert P.W.;
RT   "Structural and Functional Diversity of Glial Growth Factor Isoforms
RT   Expressed in Regenerating Peripheral Nerve and Associated Neurons.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF194993; AAG28433.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
```

```
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002154; Neuregulin.
DR   InterPro; IPR002114; PTS_HPr_ser.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
DR   PROSITE; PS005889; PTS_HPR_SER; UNKNOWN_1.
KW   Immunoglobulin domain.
FT   NON_TER        1
FT   NON_TER        1
SQ   SEQUENCE   782 AA;  86036 MW;  F6474A68F4E27BDE CRC64;

Qy     1 HFPKPCRDRDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
           || ::||| |||||  ::: |: ||||| | :| ::|||||||||||| ::
Db   316 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 363

Query Match             37.0%;  Score 102.5;  DB 11;  Length 782;
Best Local Similarity   31.2%;  Pred. No. 6.5e-06;
Matches  15;  Conservative  13;  Mismatches  19;  Indels  1;  Gaps  1;


RESULT 14
ID   Q91MZ0       PRELIMINARY;     PRT;    89 AA.
AC   Q91MZ0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   EGF-LIKE GROWTH FACTOR.
OS   lumpy skin disease virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Capripoxvirus.
OX   NCBI_TaxID=59509;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEETHLING 2490;
RX   MEDLINE=21329495; PubMed=11435593;
RA   Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "Genome of lumpy skin disease virus.";
RL   J. Virol. 75:7122-7130(2001).
SQ   SEQUENCE   89 AA;  10646 MW;  1D5F3FD7D06174E0 CRC64;

Qy     7 DKDLAYCLNDGECFVIETLTGSHK-----HCRCKEGYQGVRC 43
           ||  :||| | || : : | | :| :||:||||:||
Db    40 DKSINFCLNGGTCYKTFPILSYNKKPLMFCRCKLGYEGVRC 80

Query Match             35.7%;  Score 99;  DB 12;  Length 89;
Best Local Similarity   43.9%;  Pred. No. 2e-06;
Matches  18;  Conservative  6;  Mismatches  13;  Indels  4;  Gaps  1;


RESULT 15
001768
ID   001768       PRELIMINARY;     PRT;    2192 AA.
AC   001768;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 242.7 KDA PROTEIN.
GN   T21E3.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Du Z., Le T.T.,
RT   "The sequence of C. elegans cosmid T21E3.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
HSSP; Q07954; 1CR8.
EMBL; AF003133; AAB54138.2; -.
InterPro; IPR000561; EGF-like.
InterPro; IPR002049; Laminin_EGF.
InterPro; IPR000033; Ldl_receptor_rep.
InterPro; IPR021172; LDL_recept_A.
Pfam; PF00008; EGF; 8.
Pfam; PF00057; ldl_recept_a; 18.
Pfam; PF00058; ldl_recept_b; 8.
PRINTS; PR00011; EGFLAMININ.
PRINTS; PR00261; LDLRECEPTOR.
SMART; SM00181; EGF; 23.
SMART; SM00192; LDLa; 18.
SMART; SM00135; LY; 7.
PROSITE; PS00022; EGF_1; UNKNOWN_9.
PROSITE; PS01186; EGF_2; 7.
PROSITE; PS01209; LDLRA_1; 10.
PROSITE; PS50068; LDLRA_2; 12.
EGF-like domain; Glycoprotein; Hypothetical protein;
SQ   SEQUENCE   2192 AA;  242666 MW;  F4762A5EBCA45BDA CRC64;

Qy     5 CRDRDLAYCLNDGECFVIETLT-GSHKHCRCKEGYQGVRCDQ 45
           || || :|||| | || :: |:|| |:||||||:||
Db   1918 CDD----YCTNNSKC----TIINGTHFECDCKPGFKGLRCD 1951

Query Match             33.4%;  Score 92.5;  DB 5;  Length 2192;
Best Local Similarity   42.9%;  Pred. No. 0.00049;
Matches  18;  Conservative  8;  Mismatches  7;  Indels  9;  Gaps  3;

Search completed: November 7, 2002, 10:02:59
Job time : 27 secs
```

```
OM protein - protein search, using sw model

                    GenCore version 5.1.3
                    Copyright (c) 1993 - 2002 Compugen Ltd.

Run On:             November 7, 2002, 09:59:08 ; Search time 11 seconds
                                              (without alignments)
                                              165.438 Million cell updates/sec

Title:              US-09-480-977-4
Perfect score:      277
Sequence:           1 HFKPCRDKDLAYCLNDGECF..........SHKHCRCKEGYQGVRCDQFL 47

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched:           105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters:    105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :          SwissProt_40:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result         %
No.   Score  Query  Match  Length  DB  ID           Description
             Match

 1    277    100.0    1     713     1  035181       mus musculu
 2    277    100.0    1     720     1  NRG3_HUMAN   homo sapien
 3    126.5   45.7    1     677     1  P56975       homo sapien
 4    113.5   41.0    1     639     1  093383       xenopus lae
 5    110     39.7    1     115     1  002297       h pro-neure
 6    104.5   37.7    1     296     1  Q9wtx4       mus musculu
 7    104     37.5    1     756     1  Q15491       homo sapien
 8    102.5   37.0    1     662     1  P56974       mus musculu
 9    102     37.0    1     602     1  P43322       r pro-neure
10     98.5   35.6    1     199     1  005199       gallus gall
11     90     32.5    1     850     1  005199       gallus gall
11     90     32.5    1     868     1  NRG2_HUMAN   homo sapien
12     89.5   32.3    1     169     1  014944       rattus norv
13     85     30.7    1      80     1  P08441       shope fibro
14     83     30.0    1    2531     1  001705       mus musculu
15     82.5   29.8    1     177     1  005928       mus musculu
16     82.5   29.8    1     178     1  Q9tc5        bos taurus
17     82     29.6    1      85     1  P08072       drosophila
18     82     29.6    1     230     1  001083       myxoma viru
19     80.5   29.1    1     125     1  Q9j524       fowlpox vir
20     80.5   29.1    1     178     1  P35070       homo sapien
21     80.5   29.1    1     140     1  P01136       vaccinia vi
22     79.5   28.7    1     142     1  P20494       vaccinia vi
23     79.5   28.7    1     159     1  P48030       mus musculu
24     79.5   28.7    1     159     1  P01134       rattus norv
25     79     28.5    1    2531     1  NTC1_RAT     rattus norv
26     78     28.2    1    1207     1  EGF_HUMAN    homo sapien
27     77.5   28.0    1     714     1  P01133       rattus norv
27     77.5   28.0    1     714     1  DLL1_RAT     rattus norv
28     77.5   28.0    1     722     1  061483       mus musculu
29     76     27.4    1     484     1  LEM2_PIG     sus scrofa
30     76     27.4    1    1217     1  EGF_MOUSE    mus musculu
31     76     27.4    1    2139     1  CRB_DROME    drosophila
32     75     27.1    1     611     1  LEM2_CANFA   canis famil
33     74.5   26.9    1     603     1  FA12_CAVPO   cavia porce
34     74.5   26.9    1     723     1  DLL1_HUMAN   homo sapien
35     74.5   26.9    1    1429     1  LI12_CAEEL   caenorhabdi
36     74     26.7    1     551     1  LEM2_RABIT   oryctolagus
37     74     26.7    1     610     1  P16581       homo sapien
38     74     26.7    1    2871     1  LEM2_HUMAN   homo sapien
39     74     26.7    1    2871     1  FBN1_BOVIN   bos taurus
40     74     26.7    1    2871     1  FBN1_HUMAN   homo sapien
41     74     26.7    1    2871     1  FBN1_MOUSE   mus musculu
42     73     26.4    1     294     1  GRK_DROME    drosophila
43     73     26.4    1     409     1  MFGM_PIG     sus scrofa
44     73     26.4    1    1964     1  NTC4_MOUSE   mus musculu
45     73     26.4    1    2524     1  NOTC_XENLA   xenopus lae
```

ALIGNMENTS

```
RESULT 1
NRG3_MOUSE

ID   NRG3_MOUSE      STANDARD;      PRT;    713 AA.
AC   035181;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-3 precursor (Pro-NRG3) [Contains: Neuregulin-3 (NRG-
DE   3)].
GN   NRG3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97420720; PubMed=9275162;
RA   Zhang D., Sliwkowski M.X., Mark M., Frantz G., Akita R., Sun Y.,
RA   Hillan K., Crowley C., Brush J., Godowski P.J.;
RT   "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that
RT   binds and activates ErbB4.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).
CC   -!- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.
CC       BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR,
CC       ERBB2 OR ERBB3 RECEPTORS.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC       A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC       MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY
CC       NEURONS.
CC   -!- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS,
CC       DETECTED MAINLY IN THE NERVOUS SYSTEM. IN E16 EMBRYOS, DETECTED IN
CC       THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND
CC       SPINAL GANGLIA. IN ADULTS, EXPRESSED IN SPINAL CORD, AND NUMEROUS
CC       BRAIN REGIONS.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC       OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN (BY SIMILARITY).
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM (BY SIMILARITY).
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----------------------------------------------------------------------
```

RESULT 2

```
ID   NRG3_HUMAN     STANDARD;     PRT;     720 AA.
AC   P56975;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-3 precursor (Pro-NRG3) [Contains: Neuregulin-3 (NRG-3)].
GN   NRG3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=97420720; PubMed=9275162;
RA   Zhang D., Sliwkowski M.X., Mark M., Frantz G., Akita R., Sun Y.,
RA   Hillan K., Crowley C., Brush J., Godowski P.J.;
RT   "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that
RT   binds and activates ErbB4.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).
CC   -!- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.
CC       BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR,
CC       ERBB2 OR ERBB3 RECEPTORS.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC       A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC       MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN MOST REGIONS OF THE BRAIN
CC       WITH THE EXCEPTION OF CORPUS CALLOSUM. EXPRESSED AT LOWER LEVEL IN
CC       TESTIS. NOT DETECTED IN HEART, PLACENTA, LUNG, LIVER, SKELETAL
```

```
RX   MEDLINE=99316087; PubMed=10383827;
RA   Yang J.F., Zhou H., Choi R.C., Ip N.Y., Peng H.B., Tsim K.W.K.;
RT   "A cysteine-rich form of Xenopus neuregulin induces the expression of
RT   acetylcholine receptors in cultured myotubes.";
RL   Mol. Neurosci. 13:415-429(1999).
CC   -!- FUNCTION: DIRECT LIGAND FOR THE ERBB TYROSINE KINASE RECEPTORS.
CC       INDUCES EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC NUCLEI.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A
CC       PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
CC       BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: ISOFORMS ALPHA (SHOWN HERE) AND
CC       CRD/CRD-NRG1 (CYSTEINE RICH DOMAIN); ARE PRODUCED BY ALTERNATIVE
CC       SPLICING. ISOFORMS HAVE ALPHA OR BETA-TYPE EGF-LIKE DOMAINS.
CC   -!- TISSUE SPECIFICITY: ISOFORM ALPHA1 IS EXPRESSED IN BRAIN AND
CC       MUSCLE. ISOFORM CRD IS EXPRESSED IN BRAIN AND SPINAL CORD, BUT
CC       AT VERY LOW LEVEL IN MUSCLE.
CC   -!- DEVELOPMENTAL STAGE: STRONG EXPRESSION IN DEVELOPING BRAIN AND
CC       SPINAL CORD OF THE EMBRYO. ALSO EXPRESSED IN THE MYOTOMAL MUSCLE.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC       OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN.
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM.
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC   -------------------------------------------------------------------------
```

```
DR   EMBL; AF076618; AAC26804.1; -.
DR   EMBL; AF142632; AAD33893.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_C2.
DR   InterPro; IPR002154; Neuregulin.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Alternative splicing.
FT   CHAIN         1    259       PRO-NEUREGULIN ALPHA1 (BY SIMILARITY).
FT   CHAIN         1    677       NEUREGULIN ALPHA1 (BY SIMILARITY).
FT   DOMAIN        1    260       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    261    280       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT   DOMAIN      281    677       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       50    119       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      188    232       EGF-LIKE.
FT   DISULFID     57    116       BY SIMILARITY.
FT   DISULFID    192    206       BY SIMILARITY.
FT   DISULFID    200    220       BY SIMILARITY.
FT   DISULFID    222    231       BY SIMILARITY.
FT   DOMAIN        1     25       LYS-RICH.
FT   CARBOHYD    124    124       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC...) (POTENTIAL).
```

```
RESULT 4
ID   NRG1_HUMAN     STANDARD;      PRT;   639 AA.
AC   Q02297; Q02298; Q02299; Q12779; Q12780; Q12781; Q12782; Q12783;
AC   Q12784; Q07110; Q07111; Q9UPE3; O14667;
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1 (Neu
DE   differentiation factor) (Heregulin) (HRG) (Breast cancer cell
DE   differentiation factor p45) (Acetylcholine receptor inducing activity)
DE   (ARIA) (Sensory and motor neuron-derived factor) (Glial growth
DE   factor)].
GN   NRG1 OR HGL OR NDF OR HRGA OR GGF OR SMDF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA1; BETA2 AND BETA3), AND
RP   PARTIAL SEQUENCE.
RX   MEDLINE=92271253; PubMed=1350381;
RA   Holmes W.E., Sliwkowski M.X., Akita R.W., Henzel W.J., Lee J.,
RA   Park J.W., Yansura D., Abadi N., Raab H., Lewis G.D., Shepard H.M.,
RA   Kuang W.-J., Wood W.I., Goeddel D.V., Vandlen R.L.;
RT   "Identification of heregulin, a specific activator of p185erbB2.";
RL   Science 256:1205-1210(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA1; ALPHA2B; ALPHA3; BETA1A; BETA2
RP   AND BETA3).
RC   TISSUE=Pituitary, and Kidney adenocarcinoma;
RX   MEDLINE=94158863; PubMed=7509448;
RA   Wen D., Suggs S.V., Karunagaran D., Liu N., Cupples R.L., Luo Y.,
RA   Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,
RA   Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanigahara D.,
RA   Koski R.A., Yarden Y.;
RT   "Structural and functional aspects of the multiplicity of Neu
RT   differentiation factors.";
RL   Mol. Cell. Biol. 14:1909-1919(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX   MEDLINE=92208945; PubMed=1348215;
RA   Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,
RA   Levy R.B., Yarden Y.;
RT   "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein
RT   that induces differentiation of mammary tumor cells.";
RL   Cell 69:205-216(1992).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS BETA3 AND GGF2).
RC   TISSUE=Brain;
RX   MEDLINE=93205115; PubMed=8096067;
RA   Marchionni M.A., Goodearl A.D.J., Chen M.S., Bermingham-McDonogh O.,
```

```
Qy     1  HFKPCRDRDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ  45
          |:|| |:|||||| :|:|||||:||||||:| |:||| ||:|| |
Db   188  HLIRCSDKEKTYCVNGGECYVLNGITSSNQFMCKCKPGFTGARCTE 233
```

```
Query Match   45.7%;   Score 126.5;   DB 1;   Length 677;
Best Local Similarity   43.5%;   Pred. No. 6.6e-08;
Matches 20;   Conservative 10;   Mismatches 15;   Indels 1;   Gaps 1;
```

```
SQ   SEQUENCE   677 AA;  75794 MW;  49279E8F5BAE396F CRC64;
FT   VARSPLIC      1    136   MAEKKKKVKEGRGKGRKGKKDRKGRKRAEGSDGGAAASPKLKE
                              IKTQSVQEGKKLVLKCQAVSEQPSLKFRWFKGKEIGAKNK
                              PDSKPEHIKIRGRKKSSELQISKASSADNGEYKCMVSNQLG
                              NDTVTVNVTIVPK -> MSEDTAEGLQNQCSEQSEQSDPPSAE
                              LQNEESMPETQDEEETTHGITGLAITCCVCLEADRLRICLN
                              SEKICIPILACLISLCLCIAGLKWVFVDYEFVSPTHLD
                              PGHRGQDLILYTDTAPSTLVPSSVRRLPVIIPTTDSKAAVT
                              FKFGTSLLPTE (IN ISOFORM CRD)
FT   VARSPLIC    223    252   KPGFTGARCTETDPLRVVRSEKHLGIEFME -> PNEFTGD
FT   VARSPLIC    677    677   RCQNYVMASFYK (IN ISOFORM CRD).
```

```
RA   Kirk C., Hendricks M., Danehy F., Misumi D., Sudhalter J.,
RA   Kobayashi K., Wroblewski D., Lynch C., Baldassare M., Hiles I.,
RA   Davis J.B., Hsuan J.J., Totty N.F., Otsu M., McBurney R.N.,
RA   Waterfield M.D., Stroobant P., Gwynne D.;
RT   "Glial growth factors are alternatively spliced erbB2 ligands
RT   expressed in the nervous system.";
RL   Nature 362:312-318(1993).
RN   [5]
RP   SEQUENCE FROM N.A. OF GAMMA-HEREGULIN FUSION PROTEIN.
RC   TISSUE=Breast cancer;
RX   MEDLINE=97472144; PubMed=9333014;
RA   Schaefer G., Fitzpatrick V.D., Sliwkowski M.X.;
RT   "Gamma-heregulin: a novel heregulin isoform that is an autocrine
RT   growth factor for the human breast cancer cell line, MDA-MB-175.";
RL   Oncogene 15:1385-1394(1997).
RN   [6]
RP   SEQUENCE OF 1-210 FROM N.A.
RA   Schoumacher F., Herzer S., Flury N., Kueng W., Mueller H.,
RA   Eppenberger U.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 19-27.
RX   MEDLINE=93367731; PubMed=7689552;
RA   Culouscou J.-M., Plowman G.D., Carlton G.W., Green J.M., Shoyab M.;
RT   "Characterization of a breast cancer cell differentiation factor that
RT   specifically activates the HER4/p180erbB4 receptor.";
RL   J. Biol. Chem. 268:18407-18410(1993).
RN   [8]
RP   CHROMOSOMAL TRANSLOCATION.
RX   MEDLINE=99455251; PubMed=10523851;
RA   Wang X.-Z., Jolicoeur E.M., Conte N., Chaffanet M., Zhang Y.,
RA   Mozziconacci M.-J., Feiner H., Birnbaum D., Pebusque M.-J., Ron D.;
RT   "Gamma-heregulin is the product of a chromosomal translocation fusing
RT   the DOC4 and HGL/NRG1 genes in the MDA-MB-175 breast cancer cell
RT   line.";
RL   Oncogene 18:5718-5721(1999).
RN   [9]
RP   CHROMOSOMAL TRANSLOCATION.
RX   MEDLINE=20055180; PubMed=10597312;
RA   Liu X., Baker E., Eyre H.J., Sutherland G.R., Zhou M.;
RT   "Gamma-heregulin: a fusion gene of DOC-4 and neuregulin-1 derived from
RT   a chromosome translocation.";
RL   Oncogene 18:7110-7114(1999).
RN   [10]
RP   STRUCTURE BY NMR OF 175-241 (ISOFORM ALPHA).
RX   MEDLINE=94341264; PubMed=8062828;
RA   Nagata K., Kohda D., Hatanaka H., Ichikawa S., Matsuda S.,
RA   Yamamoto T., Suzuki A., Inagaki F.;
RT   "Solution structure of the epidermal growth factor-like domain of
RT   heregulin-alpha, a ligand for p180erbB-4.";
RL   EMBO J. 13:3517-3523(1994).

CC   -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC       RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC       RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM
CC       DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF
CC       EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING
CC       EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESSICLES DURING
CC       THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING
CC       LOBULOALVEOLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND
CC       AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING
CC       SCHWANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE
CC       MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART.
CC   -!- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
CC       REGION OF LIMK1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC       A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC       MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE. THE SECRETED
CC       ISOFORM GGF2 HAS A SIGNAL PEPTIDE. THE ISOFORM BETA3 MAY BE
CC       NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS; ALPHA (SHOWN HERE);
CC       ALPHA1, ALPHA2B, ALPHA3, BETA1, BETA1A, BETA2, BETA3/GGFHEB1,
CC       GGF2/GGFHPP2 AND SMDF (AC Q15491); ARE PRODUCED BY ALTERNATIVE

CC       SPLICING. THEY HAVE BEEN CLASSIFIED AS TYPE I NRGS (VARIANTS WITH
CC       AN IG DOMAIN AND A GLYCOSYLATION DOMAIN; ALPHA AND BETA), TYPE II
CC       NRGS (VARIANTS WITH AN IG DOMAIN BUT NO GLYCOSYLATION DOMAIN;
CC       GGF2) AND TYPE III NRGS (VARIANTS WITH A CYS-RICH DOMAIN; SMDF).
CC       ALL THESE ISOFORMS PERFORM DISTINCT TISSUE-SPECIFIC FUNCTIONS.
CC   -!- TISSUE SPECIFICITY: TYPE I ISOFORMS ARE THE PREDOMINANT FORMS
CC       EXPRESSED IN THE ENDOCARDIUM. ISOFORM ALPHA IS EXPRESSED IN
CC       BREAST, OVARY, TESTIS, PROSTATE, HEART, SKELETAL MUSCLE, LUNG,
CC       PLACENTA, LIVER, KIDNEY, SALIVARY GLAND, SMALL INTESTINE AND BRAIN,
CC       BUT NOT IN UTERUS, STOMACH, PANCREAS, AND SPLEEN. ALPHA2 IS THE
CC       PREDOMINANT FORM IN MESENCHYMAL CELLS AND IN NONNEURONAL ORGANS,
CC       WHEREAS BETA1 IS THE MAJOR NEURONAL FORM. BETA3 IS EXPRESSED IN
CC       SPINAL CORD AND BRAIN. GGF2 IS THE MAJOR FORM IN SKELETAL MUSCLE
CC       CELLS; IN THE NERVOUS SYSTEM IT IS EXPRESSED IN SPINAL CORD AND
CC       BRAIN. ALSO DETECTED IN ADULT HEART, PLACENTA, LUNG, LIVER,
CC       KIDNEY, AND PANCREAS.
CC   -!- DEVELOPMENTAL STAGE: DETECTABLE AT EARLY EMBRYONIC AGES.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC       OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN.
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM.
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC       SIMILARITY)
CC   -!- DISEASE: INVOLVED IN A RARE T(8;11) CHROMOSOMAL TRANSLOCATION THAT
CC       FUSES THE 5'END OF ODZ4 TO NRG1 (ISOFORM BETA3). THE PRODUCT OF
CC       THIS TRANSLOCATION HAS FIRST BE THOUGHT TO BE AN ALTERNATIVE
CC       SPLICED ISOFORM, CALLED GAMMA-HEREGULIN. GAMMA-HEREGULIN IS A
CC       SOLUBLE ACTIVATING LIGAND FOR THE ERBB2-ERBB3 RECEPTOR COMPLEX AND
CC       ACTS AS AN AUTOCRINE GROWTH FACTOR IN A SPECIFIC BREAST CANCER
CC       CELL LINE (MDA-MB-175). NOT DETECTED IN BREAST CARCINOMA SAMPLES,
CC       INCLUDING DUCTAL, LOBULAR, MEDULLARY, AND MUCINOUS HISTOLOGICAL
CC       TYPES, NEITHER IN OTHER BREAST CANCER CELL LINES.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
```

```
DR   EMBL; M94165; AAA58638.1; -.
DR   EMBL; M94166; AAA58639.1; -.
DR   EMBL; M94167; AAA58640.1; -.
DR   EMBL; M94168; AAA58641.1; -.
DR   EMBL; L12261; AAB59358.1; -.
DR   EMBL; U02325; AAA19950.1; -.
DR   EMBL; U02326; AAA19951.1; -.
DR   EMBL; U02327; AAA19952.1; -.
DR   EMBL; U02328; AAA19953.1; -.
DR   EMBL; U02329; AAA19954.1; -.
DR   EMBL; U02330; AAA19955.1; -.
DR   EMBL; L12260; AAD01795.1; -.
DR   EMBL; AF026146; AAD01795.1; -.
DR   EMBL; AF009227; AAC51756.1; ALT_INIT.
DR   PDB; 1HRE; 15-OCT-94.
DR   PDB; 1HRF; 15-OCT-94.
DR   MIM; 142445; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002154; Neuregulin.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   PROSITE; PS00022; EGF_1; 1.
```

DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Multigene family; Polymorphism; 3D-structure.
KW   Alternative splicing; Chromosomal translocation.

Query Match
Best Local Similarity   41.0%; Score 113.5; DB 1; Length 639;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

```
Qy     1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ  45
         |:||:|||| |||||||||:|:||: | ||: |||||:| ||| |:||
Db   177 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 222
```

RESULT 5

```
ID   NRG4_MOUSE     STANDARD;      PRT;   115 AA.
AC   Q9WTX4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Pro-neuregulin-4, short isoform (pro-NRG4) (Contains: Neuregulin-4
DE   (NRG-4)).
GN   NRG4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=99276098; PubMed=10348342;
RA   Harari D., Tzahar E., Romano J., Shelly M., Pierce J.H., Andrews G.C.,
RA   Yarden Y.;
RT   "Neuregulin-4: a novel growth factor that acts through the ErbB-4
RT   receptor tyrosine kinase.";
RL   Oncogene 18:2681-2689(1999).
CC   -!- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE
CC       RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC       RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1,
CC       ERBB2 AND ERBB3 RECEPTORS.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC       A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC       MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEAKLY EXPRESSED
CC       IN MUSCLE.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC       OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN (BY SIMILARITY).
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM (BY SIMILARITY).
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
```

```
DR   EMBL; AF083067; AAD21874.1; -.
DR   MGD; MGI:1933833; Nrg4.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001336; EGF_1.
DR   Pfam; PF00008; EGF; 1.
DR   PRINTS; PR00009; EGFGF
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   Growth factor; EGF-like domain; Glycoprotein; Transmembrane;
KW   Multigene family; Alternative splicing.
FT   CHAIN         1    115       PRO-NEUREGULIN-4, MEMBRANE-BOUND FORM.
FT   CHAIN         1     61       NEUREGULIN-4.
FT   CHAIN         1     62       PRO-NEUREGULIN-4.
FT   DOMAIN        1     83       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     63     83       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT   DOMAIN       84    115       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        5     46       EGF-LIKE.
FT   DISULFID      9     23       BY SIMILARITY.
FT   DISULFID     17     34       BY SIMILARITY.
FT   DISULFID     36     45       BY SIMILARITY.
FT   CARBOHYD     39     39       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   115 AA;  12743 MW;  989A1E376F85789 CRC64;
```

Query Match
Best Local Similarity   42.2%; Score 110; DB 1; Length 115;
Matches 19; Conservative 8; Mismatches 16; Indels 2; Gaps 1;

```
Qy     1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQ  45
         |:||| |||  |||| | ||: | |:|: ||: |||||: ||| ||:
Db     5 HEQPCGPRHRSFCLNGGGICYVIPIPS--PFCRCIENYTGARCEE  47
```

RESULT 6

```
ID   SMDF_HUMAN     STANDARD;      PRT;   296 AA.
AC   Q15491;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neuregulin-1, sensory and motor neuron-derived factor isoform.
GN   NRG1 OR HGL OR NDF OR HRGA OR GGF OR SMDF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain stem, and Cerebellum;
RX   MEDLINE=95301541; PubMed=7782315;
RA   Ho W.-H., Armanini M.P., Nuijens A., Phillips H.S., Osheroff P.L.;
RT   "Sensory and motor neuron-derived factor. A novel heregulin variant
RT   highly expressed in sensory and motor neurons.";
RL   J. Biol. Chem. 270:14523-14532(1995).
CC   -!- FUNCTION: THE ISOFORM SMDF MAY PLAY A ROLE IN MOTOR AND SENSORY
CC       NEURON DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: SECRETED. MAY POSSESS AN INTERNAL UNCLEAVED
CC       SIGNAL SEQUENCE.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS OF NRG1 ARE PRODUCED BY
CC       ALTERNATIVE SPLICING. EXCEPT FOR SMDF THEY ARE IN ENTRY AC Q02297.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD MOTOR
CC       NEURONS, DORSAL ROOT GANGLION NEURONS, AND BRAIN. PREDOMINANT
CC       ISOFORM EXPRESSED IN SENSORY AND MOTOR NEURONS. NOT DETECTED IN
CC       ADULT HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, AND
CC       PANCREAS. NOT EXPRESSED IN FETAL LUNG, LIVER, AND KIDNEY.
CC   -!- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN DEVELOPING SPINAL MOTOR
CC       NEURONS AND IN DEVELOPING CRANIAL NERVE NUCLEI. EXPRESSION IS
CC       MAINTAINED ONLY IN BOTH ADULT MOTOR NEURONS AND DORSAL ROOT
CC       GANGLION NEURONS.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L41827; AAC41764.1; -.
DR   HSSP; Q12784; 1HRE.
MIM; 142445; -.
DR   InterPro; IPR000561; EGF-like.
DR   Pfam; PF00008; EGF; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   Growth factor; EGF-like domain; Transmembrane; Multigene family;
KW   Alternative splicing.
FT   TRANSMEM     76     100       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT   DOMAIN       58      91       CYS-RICH.
FT   DOMAIN      211     232       SER/THR-RICH.
FT   DOMAIN      233     277       EGF-LIKE.
FT   DISULFID    237     251       BY SIMILARITY.
FT   DISULFID    245     265       BY SIMILARITY.
FT   DISULFID    267     276       BY SIMILARITY.
SQ   SEQUENCE   296 AA;  31685 MW;  8D417432177EB02 CRC64;

Qy      1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL    47
           :||: :||||||||||||||||||| ||||| ||||||||||||||||||
Db    233  HLVKCAEKEKTPCVNGGECFMVKDLSNPSRYLCKCPNEFTGDRCQNTV    280
```

RESULT 7
```
NRG2_MOUSE

ID   NRG2_MOUSE     STANDARD;      PRT;   756 AA.
AC   P56974;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-2 precursor (Pro-NRG2) (Contains: Neuregulin-2 (NRG-2)
DE   (Divergent of neuregulin 1) (DON-1)].
GN   NRG2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS NRG2-5; NRG2-10 AND NRG2-16A).
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=97311398; PubMed=9168115;
RA   Carraway K.L. III, Weber J.L., Unger M.J., Ledesma J., Yu N.,
RA   Gassmann M., Lai C.;
RT   "Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine
RT   kinases.";
RL   Nature 387;512-516(1997).
RN   [2]
RP   SEQUENCE OF 150-756 FROM N.A. (ISOFORMS DON-1M AND DON-1S).
RC   TISSUE=Choroid plexus;
RX   MEDLINE=97342638; PubMed=9199335;
RA   Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,
RA   Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,
RA   Gearing D.P.;
RT   "Characterization of a neuregulin-related gene, Don-1, that is highly
RT   expressed in restricted regions of the cerebellum and hippocampus.";
RL   Mol. Cell. Biol. 17:4007-4014(1997).
-!-  FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
     RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
     RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
     ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
     HETERODIMERIZATION WITH THE EGF RECEPTOR.
-!-  SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
     A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
     MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
-!-  ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; DON-1M, DON-1S/NRG2-5,
     NRG2-10 AND NRG2-16A (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
     SPLICING.
-!-  TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THE BRAIN, WITH LOWER
     LEVELS IN THE LUNG. IN THE CEREBELLUM, FOUND IN GRANULE AND
     PURKINJE CELLS.
-!-  DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
     OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
     PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
     DIMERIZATION (BY SIMILARITY).
-!-  DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
     DOMAIN (BY SIMILARITY).
-!-  PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
     EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
     FORM (BY SIMILARITY).
-!-  PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
     SIMILARITY).
DR   Pfam; PF00047; ig; 1.
SMART; SM00408; IGc2; 1.
DR   SMART; SM00181; EGF; 1.
-!-  SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-!-  SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
-!-  SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
MGD; MGI:1098246; Nrg2.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR002154; Neuregulin.
Pfam; PF02158; Neuregulin; 2.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Multigene family; Alternative splicing.
FT   PROPEP        1      19       BY SIMILARITY.
FT   CHAIN        20     756       PRO-NEUREGULIN-2, MEMBRANE-BOUND FORM.
FT   CHAIN        20     314       NEUREGULIN-2.
FT   DOMAIN       20     315       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    316     336       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT   DOMAIN      337     756       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      158     226       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      238     248       SER/THR-RICH.
FT   DOMAIN      249     290       EGF-LIKE.
FT   DOMAIN      627     633       POLY-PRO.
FT   DISULFID    165     219       BY SIMILARITY.
FT   DISULFID    253     267       BY SIMILARITY.
FT   DISULFID    261     278       BY SIMILARITY.
FT   DISULFID    280     289       BY SIMILARITY.
FT   CARBOHYD     55      55       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    186     186       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    254     254       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    296     296       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   VARSPLIC    280     280       C -> G (IN ISOFORM NRG2-10).
FT   VARSPLIC    281     756       MISSING (IN ISOFORM NRG2-10).
FT   VARSPLIC    282     330       VGYTGDRCQQFAMVNFSKHLGFELKEAEELYQRVLTIPGI
                                   CVALLVVG -> NGFFGQRCLEKLPLRLYMPDPRQSVLWDT
                                   PGTGVSSSQNWSTSPSTLDLN (IN ISOFORM DON-1S).
FT   VARSPLIC    331     756       MISSING (IN ISOFORM DON-1S).
FT   VARSPLIC    282     307       VGYTGDRCQQFAMVNFSKHLGFELKE -> NGFFGQRCLEK
                                   LPLRLYMPDPKQK (IN ISOFORM DON-1M).
SQ   SEQUENCE   756 AA;  82213 MW;  51D85DC518B5678E CRC64;

Qy      1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQF    46
           :||: :||||||||||||||||||| |||||||||||||||||||||
Db    249  HARCNEFAKSYCVNGGVCYYIEGI--NQLSCKCPVGYTGDRCQQF    292
```

Query Match   37.5%;   Score 104;   DB 1;   Length 756;
Best Local Similarity   39.1%;   Pred. No. 3.7e-05;
Matches   18;   Conservative   8;   Mismatches   18;   Indels   2;   Gaps   1;

RESULT 8
NRG1_RAT

(additional left-column fragments)

Query Match   37.7%;   Score 104.5;   DB 1;   Length 296;
Best Local Similarity   31.2%;   Pred. No. 1.3e-05;
Matches   15;   Conservative   14;   Mismatches   18;   Indels   1;   Gaps   1;

```
ID   NRG1_RAT       STANDARD;      PRT;   662 AA.
AC   P43322; P43323; P43324; P43325; P43326; P43327; P43328;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-1 precursor (Pro-NRG1) (Contains: Neuregulin-1 (Neu
DE   differentiation factor) (Heregulin) (HRG) (Acetylcholine receptor
DE   inducing activity) (ARIA) (Sensory and motor neuron-derived factor)
DE   (Glial growth factor)].
GN   NRG1 OR NDF.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RC   TISSUE=Fibroblast;
RX   MEDLINE=94158863; PubMed=7509448;
RA   Wen D., Suggs S.V., Karunagaran D., Liu N., Cupples R.L., Luo Y.,
RA   Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,
RA   Meng S., Lu H.S., Hu S., Chang D., Yang W., Yanigahara D.,
RA   Koski R.A., Yarden Y.;
RT   "Structural and functional aspects of the multiplicity of Neu
RT   differentiation factors.";
RL   Mol. Cell. Biol. 14:1909-1919(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA2C/NDF44), AND PARTIAL SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE=92200945; PubMed=1348215;
RA   Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D.,
RA   Levy R.B., Yarden Y.;
RT   "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein
RT   that induces differentiation of mammary tumor cells.";
RL   Cell 69:205-216(1992).
RN   [3]
RP   SEQUENCE OF 14-36.
RX   MEDLINE=92208945; PubMed=9685409;
RA   Trail G., Hu S., Silbiger S.M., Levy R.B., Koski R.A., Lu H.S.,
RA   Yarden Y.;
RT   "Neu differentiation factor: a transmembrane glycoprotein containing
RT   an EGF domain and an immunoglobulin homology unit.";
RL   Cell 69:559-572(1992).
RN   [4]
RP   REGULATION OF PROCESSING (ISOFORM ALPHA2C/NDF44).
RX   MEDLINE=99069430; PubMed=9852099;
RA   Liu X., Hwang H., Cao L., Wen D., Liu N., Graham R.M., Zhou M.;
RT   "Release of the neuregulin functional polypeptide requires its
RT   cytoplasmic tail.";
RL   J. Biol. Chem. 273:34335-34340(1998).
RN   [5]
RP   INTERACTION WITH LIMK1.
RX   MEDLINE=98352096; PubMed=9685409;
RA   Wang J.Y., Frenzel K.E., Wen D., Falls D.L.;
RT   "Transmembrane neuregulins interact with LIM kinase 1, a cytoplasmic
RT   protein kinase implicated in development of visuospatial cognition.";
RL   J. Biol. Chem. 273:20525-20534(1998).
CC   -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC       RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC       RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM
CC       DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF
CC       EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING
CC       EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING
CC       THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING
CC       LOBULOALVEOLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND
CC       AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING
CC       SCHWANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE
CC       MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
CC       REGION OF LIMK1.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A
CC       PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
CC       BOUND FORM DOES NOT SEEM TO BE ACTIVE.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; ALPHA2A/NDF38,
CC       ALPHA2B/NDF19, ALPHA2C/NDF44, BETA1, BETA1A, BETA2A/NDF40,
CC       BETA3/NDF4 AND BETA4/NDF42A (SHOWN HERE); ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. MOST TISSUES CONTAIN ALPHA2A
CC       AND ALPHA2B ISOFORMS. ALPHA2 AND BETA2 ARE THE PREDOMINANT FORMS
CC       IN MESENCHYMAL AND NONNEURONAL ORGANS. BETA1 IS ENRICHED IN BRAIN
CC       AND SPINAL CORD, BUT NOT IN MUSCLE AND HEART. ALPHA2C IS HIGHLY
CC       EXPRESSED IN SPINAL CORD, MODERATELY IN LUNG, BRAIN, OVARY, AND
CC       STOMACH, IN LOW AMOUNTS IN THE KIDNEY, SKIN AND HEART AND NOT
CC       DETECTED IN THE LIVER, SPLEEN, AND PLACENTA.
CC   -!- DOMAIN: THE CYTOPASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF
CC       TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION.
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN.
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM.
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U02315; AAA19940.1; -.
DR   EMBL; U02316; AAA19941.1; -.
DR   EMBL; U02317; AAA19942.1; -.
DR   EMBL; U02318; AAA19943.1; -.
DR   EMBL; U02319; AAA19944.1; -.
DR   EMBL; U02320; AAA19945.1; -.
DR   EMBL; U02321; AAA19946.1; -.
DR   EMBL; U02322; AAA19947.1; -.
DR   EMBL; U02323; AAA19948.1; -.
DR   EMBL; U02324; AAA19949.1; -.
DR   EMBL; M92430; -; NOT_ANNOTATED_CDS.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR002154; Neuregulin.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; EGF; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Multigene family; Alternative splicing.
FT   PROPEP        1     13
FT   CHAIN        14    662       PRO-NEUREGULIN-1, MEMBRANE-BOUND FORM.
FT   CHAIN        14    264       NEUREGULIN-1.
FT   DOMAIN       14    265       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    266    288       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT   DOMAIN      289    662       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       50    119       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      165    177       SER/THR-RICH.
FT   DOMAIN      178    222       EGF-LIKE.
FT   DISULFID     57    112       BY SIMILARITY.
FT   DISULFID    182    196       BY SIMILARITY.
FT   DISULFID    190    210
```

RESULT 9

```
ID   NRG1_CHICK     STANDARD;      PRT;   602 AA.
AC   Q05199; O73750; O73751; O73752;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-1 precursor (Pro-NRG1) (Contains: Neuregulin-1
DE   (Acetylcholine receptor inducing activity) (ARIA)).
GN   NRG1 OR ARIA.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS BETA1A), AND PARTIAL SEQUENCE.
RC   TISSUE=Brain, and Spinal cord;
RX   MEDLINE=93201602; PubMed=8453670;
RA   Falls D.L., Rosen K.M., Corfas G., Lane W.S., Fischbach G.D.;
RT   "ARIA, a protein that stimulates acetylcholine receptor synthesis, is
RT   a member of the neu ligand family.";
RL   Cell 72:801-815(1993).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS BETA1A; BETA2A AND BETA2B).
RC   TISSUE=Brain, and Spinal cord;
RX   MEDLINE=98150951; PubMed=9491987;
RA   Yang X., Kuo Y., Devay P., Yu C., Role L.;
RT   "A cysteine-rich isoform of neuregulin controls the level of
RT   expression of neuronal nicotinic acetylcholine receptor channels during
RT   synaptogenesis.";
RL   Neuron 20:255-270(1998).
CC   -!- FUNCTION: DIRECT LIGAND FOR THE ERBB TYROSINE KINASE RECEPTORS.
CC       THE MULTIPLE ISOFORMS PERFORM DIVERSE FUNCTIONS: CYSTEIN-RICH
CC       DOMAIN CONTAINING ISOFORMS (CRD-NRG) PROBABLY REGULATE THE
CC       EXPRESSION OF NICOTINIC ACETYLCHOLINE RECEPTORS AT DEVELOPING
CC       INTERNEURONAL SYNAPSES. THE IG-NRG ISOFORM IS REQUIRED FOR THE
CC       INITIAL INDUCTION AND/OR MAINTENANCE OF THE MATURE LEVELS OF
CC       ACETYLCHOLINE RECEPTORS AT NEUROMUSCULAR SYNAPSES.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A
CC       PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
CC       BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; ARIA/IG-NRG (SHOWN
CC       HERE), CRD-NRG-BETA1A, CRD-NRG-BETA2A AND CRD-NRG-BETA2B; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING. ARIA CONTAINS AN IG-LIKE DOMAIN,
CC       WHEREAS IN THE CRD-NRG (OR NARIA) ISOFORMS, THE EGF-LIKE DOMAIN IS
CC       REPLACED BY A CYSTEINE-RICH DOMAIN (CRD).
CC   -!- DEVELOPMENTAL STAGE: CRD-NRG ISOFORM IS DETECTED AT EMBRYONIC DAY
CC       4 (ED4) IN BOTH VISCERAL AND SOMATIC MOTOR NEURONS OF SPINAL CORD
CC       AND IS HIGHEST AT ED6. IG-NRG ISOFORM IS NOT EXPRESSED UNTIL ED 6
CC       IN SPINAL CORD. AT ED 11 BOTH ISOFORMS DISPLAY COMPARABLE LEVELS.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC       OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN.
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM.
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L11264; AAA49037.1; -.
DR   EMBL; AF045654; AAC05670.1; -.
DR   EMBL; AF045655; AAC05671.1; -.
DR   EMBL; AF045656; AAC05672.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_C2.
DR   InterPro; IPR002154; Neuregulin.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF02158; Neuregulin; 1.
DR   PRINTS; PR01089; NEUREGULIN.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Alternative splicing.
FT   CHAIN         1    602       PRO-NEUREGULIN-1, MEMBRANE-BOUND FORM.
FT   CHAIN         1    205       NEUREGULIN-1.
FT   DOMAIN        1    206       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    207    229       POTENTIAL.
FT   DOMAIN      230    602       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       42    112       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      125    136       SER/THR-RICH.
FT   DOMAIN      137    181       EGF-LIKE.
FT   DISULFID     49    105       BY SIMILARITY.
FT   DISULFID    141    155       BY SIMILARITY.
FT   DISULFID    149    169       BY SIMILARITY.
FT   DISULFID    171    180       BY SIMILARITY.
FT   CARBOHYD     21     21       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC      1    127       DTNA -> MSEVGTETFPSPSAQLSPDASLGGLPAEENMPG
FT                                KOKXYSELHIYRATLADAGEYACRVSSKLGNDSTKASVIIT
FT                                GORLVLRCETTSEYPALRPKWLKNGRETITKNRNRPENVIKPR
FT                                MWAISEGPLQYSLAPTQTDVNSSYNTVPPRLKEMKNQEVAV
FT                                (IN ISOFORM ALPHA2B).
FT   VARSPLIC    126    126       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    164    164       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    120    120       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID    212    221       BY SIMILARITY.
FT   CARBOHYD    120    120       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    164    164       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC    213    256       PNEFTGDRCQNTYVNAASFYMTSRRKRQETEKPLERKLDHSLV
FT                                KES -> QPGFTGARCTENVPMRKVQTQE (IN ISOFORM
FT                                ALPHA2A, ISOFORM ALPHA2B AND ISOFORM
FT                                ALPHA2C).
FT   VARSPLIC    231    257       MTSRRKRQETEKPLERKLDHSLVRESK -> KHLGIEPME
FT                                (IN ISOFORM BETA1).
FT   VARSPLIC    231    241       MTSRRKRQETE -> STSTPFLSLPE (IN ISOFORM
FT                                BETA3).
FT   VARSPLIC    242    662       MISSING (IN ISOFORM BETA3).
FT   VARSPLIC    231    256       MISSING (IN ISOFORM BETA2 AND ISOFORM
FT                                BETA2A).
FT   VARSPLIC    325    330       PPENVQ -> RVRTRG (IN ISOFORM ALPHA2C).
FT   VARSPLIC    446    662       MISSING (IN ISOFORM ALPHA2C).
FT   VARSPLIC    446    484       YVSAMTTPARMSFVDFHFPSSPRSPPSPEMSPPVSSMTYS
FT                                -> HNLIAELRRNKAYRSKCMQIOLSATHLRPSSITHLGFI
FT                                L (IN ISOFORM ALPHA2B).
FT   CONFLICT    485    662       MISSING (IN ISOFORM ALPHA2B).
FT   CONFLICT     90     90       K -> N (IN REF. 2).
FT   CONFLICT    137    137       T -> I (IN REF. 2).
FT   CONFLICT    208    208       Y -> S (IN REF. 2; AA SEQUENCE).
SQ   SEQUENCE   662 AA;  73288 MW;  1C31ABCF2A8EBID5 CRC64;
```

Query Match   37.0%;   Score 102.5; DB 1; Length 662;
Best Local Similarity   31.2%;   Pred. No. 4.9e-05;
Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

Qy   1 HFRKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
         | |||| | ||| || | | | | || | || | :: ::
Db 178 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 225

FT   VARSPLIC   191   198   MISSING (IN ISOFORM BETA1A, ISOFORM
FT   BETA2A AND ISOFORM BETA2B).
FT   VARSPLIC   388   405   VSAMTPARNSPVDFHTP -> HTPPTSILLAGKVSLRVS
FT   (IN ISOFORM BETA2B).
FT   VARSPLIC   406   602   MISSING (IN ISOFORM BETA2A AND ISOFORM
FT   BETA2B).
SQ   SEQUENCE   602 AA; 67453 MW; 4183C0E56CCE5D346 CRC64;
     PHREDSRVPGVAGLASTCCVCLEAERLKGCLNSEKICTAPI
     LACLLSLCLCIAGLKWVFWVDKIEEYDSPTHLDPGRIGQDPR
     STVDPTALSAWPSEVVASPFPPSLESRAEVTVQTDSSLV
     PSRPELQPSLYNRLLDVGLWSSANPSLSPPSSLEPTTASQAQ
     ATERNLQTAPKLS (IN ISOFORM BETA1A, ISOFORM
     BETA2A AND ISOFORM BETA2B).

Query Match   98.5; DB 1; Length 602;
Best Local Similarity   35.6%;
Matches 16; Conservative 11; Pred. No. 0.00014; Mismatches 20; Indels 1; Gaps 1;

Qy   1 HFRPCRDKLAYCLNDSGECFVIETLTCGSHKH-CRCKEGYQGVRCDQFL 47
           |.|:| |:|:| |||.:||| |:|: | ||| | : :| | ::
Db   137 HLTKCDIKQKAFCVNGGECYMVKDLPNPPRYLCRCPNEFTGDRCQNYV 184

RESULT 10
NRG2_HUMAN

ID   NRG2_HUMAN   STANDARD;   PRT;   850 AA.
AC   O14511;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)
DE   (Neural- and thymus-derived activator for ERBB kinases) (NTAK)
DE   (Divergent of neuregulin 1) (DON-1)].
GN   NRG2 OR NTAK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS DON-1B AND DON-1R).
RC   TISSUE=Neuroblastoma;
RX   MEDLINE=98006324; PubMed=9348101;
RA   Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,
RA   Nakagawa T., Miyagawa J., Matsushita N., Nagatsu T., Taniguchi N.,
RA   Ishiguro H.;
RT   "A novel brain-derived member of the epidermal growth factor family
RT   that interacts with ErbB3 and ErbB4.";
RL   J. Biochem. 122:675-680(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).
RC   TISSUE=Lung, and Fetal brain;
RX   MEDLINE=99295836; PubMed=10369162;
RA   Ring H.Z., Chang H., Guilbot A., Brice A., LeGuern E., Francke U.;
RT   "The human neuregulin 2 (NRG2) gene: cloning, mapping and evaluation
RT   as a candidate for the autosomal recessive form of Charcot-Marie-Tooth
RT   disease linked to 5q.";
RL   Hum. Genet. 104:326-332(1999).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1).
RC   TISSUE=Fetal brain;
RX   MEDLINE=97342638; PubMed=9199335;
RA   Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,
RA   Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,
RA   Gearing D.P.;
RT   "Characterization of a neuregulin-related gene, Don-1, that is highly
RT   expressed in restricted regions of the cerebellum and hippocampus.";
RL   Mol. Cell. Biol. 17:4007-4014(1997).
CC   -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC   RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC   RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC   ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
CC   HETERODIMERIZATION WITH THE EGF RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC   A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC   MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: 8 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4, 5, 6,
CC   DON-1B AND DON-1R; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO THE CEREBELLUM IN THE ADULT.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC   OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC   PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC   DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC   DOMAIN (BY SIMILARITY).
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC   EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC   FORM (BY SIMILARITY).
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC   SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

DR   EMBL; AB005060; BAA23417.1; -.
DR   EMBL; AF119162; AAF28848.1; -.
DR   EMBL; AF119151; AAF28848.1; JOINED.
DR   EMBL; AF119152; AAF28848.1; JOINED.
DR   EMBL; AF119153; AAF28848.1; JOINED.
DR   EMBL; AF119154; AAF28848.1; JOINED.
DR   EMBL; AF119155; AAF28848.1; JOINED.
DR   EMBL; AF119156; AAF28848.1; JOINED.
DR   EMBL; AF119157; AAF28848.1; JOINED.
DR   EMBL; AF119158; AAF28848.1; JOINED.
DR   EMBL; AF119159; AAF28848.1; JOINED.
DR   EMBL; AF119160; AAF28848.1; JOINED.
DR   EMBL; AF119161; AAF28848.1; JOINED.
DR   EMBL; AF119161; AAF28849.1; -.
DR   EMBL; AF119151; AAF28849.1; JOINED.
DR   EMBL; AF119152; AAF28849.1; JOINED.
DR   EMBL; AF119153; AAF28849.1; JOINED.
DR   EMBL; AF119154; AAF28849.1; JOINED.
DR   EMBL; AF119155; AAF28849.1; JOINED.
DR   EMBL; AF119156; AAF28849.1; JOINED.
DR   EMBL; AF119157; AAF28849.1; JOINED.
DR   EMBL; AF119158; AAF28849.1; JOINED.
DR   EMBL; AF119159; AAF28849.1; JOINED.
DR   EMBL; AF119160; AAF28849.1; JOINED.
DR   EMBL; AF119160; AAF28850.1; -.
DR   EMBL; AF119151; AAF28850.1; JOINED.
DR   EMBL; AF119152; AAF28850.1; JOINED.
DR   EMBL; AF119153; AAF28850.1; JOINED.
DR   EMBL; AF119154; AAF28850.1; JOINED.
DR   EMBL; AF119155; AAF28850.1; JOINED.
DR   EMBL; AF119156; AAF28850.1; JOINED.
DR   EMBL; AF119157; AAF28850.1; JOINED.
DR   EMBL; AF119158; AAF28850.1; JOINED.
DR   EMBL; AF119159; AAF28850.1; JOINED.
DR   EMBL; AF119160; AAF28850.1; JOINED.
DR   EMBL; AF119162; AAF28850.1; JOINED.
DR   EMBL; AF119161; AAF28850.1; JOINED.
DR   EMBL; AF119159; AAF28851.1; -.
DR   EMBL; AF119151; AAF28851.1; JOINED.
DR   EMBL; AF119152; AAF28851.1; JOINED.
DR   EMBL; AF119153; AAF28851.1; JOINED.
DR   EMBL; AF119154; AAF28851.1; JOINED.
DR   EMBL; AF119155; AAF28851.1; JOINED.
DR   EMBL; AF119156; AAF28851.1; JOINED.
DR   EMBL; AF119157; AAF28851.1; JOINED.
DR   EMBL; AF119158; AAF28851.1; JOINED.
DR   EMBL; AF119160; AAF28851.1; JOINED.
DR   EMBL; AF119161; AAF28851.1; JOINED.
DR   EMBL; AF119152; AAF28852.1; -.
DR   EMBL; AF119151; AAF28852.1; JOINED.
DR   EMBL; AF119158; AAF28852.1; JOINED.
DR   EMBL; AF119161; AAF28852.1; JOINED.
DR   EMBL; AF119159; AAF28852.1; JOINED.
DR   EMBL; AF119160; AAF28852.1; JOINED.
DR   EMBL; AF119152; AAF28852.1; JOINED.

```
DR   EMBL; AF119153; AAF28852.1; JOINED.
DR   EMBL; AF119154; AAF28852.1; JOINED.
DR   EMBL; AF119155; AAF28852.1; JOINED.
DR   EMBL; AF119156; AAF28852.1; JOINED.
DR   EMBL; AF119157; AAF28852.1; JOINED.
DR   EMBL; AF119151; AAF28853.1; JOINED.
DR   EMBL; AF119152; AAF28853.1; JOINED.
DR   EMBL; AF119153; AAF28853.1; JOINED.
DR   EMBL; AF119154; AAF28853.1; JOINED.
DR   EMBL; AF119155; AAF28853.1; JOINED.
DR   EMBL; AF119156; AAF28853.1; JOINED.
DR   HSSP; Q12784; 1HRE.
DR   MIM; 603818; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_C2.
DR   InterPro; IPR002154; Neuregulin.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF02158; Neuregulin; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00408; IG2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Multigene family; Alternative splicing.
FT   PROPEP        1    111       BY SIMILARITY.
FT   CHAIN       112    850       PRO-NEUREGULIN-2, MEMBRANE-BOUND FORM.
FT   CHAIN       112    404       NEUREGULIN-2.
FT   DOMAIN      112    405       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    406    426       INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT   DOMAIN      427    850       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      250    318       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      330    340       SER/THR-RICH.
FT   DOMAIN      341    382       EGF-LIKE.
FT   DOMAIN       10     13       POLY-PRO.
FT   DOMAIN       20     30       POLY-SER.
FT   DOMAIN       33     47       POLY-SER.
FT   DOMAIN       87     90       POLY-ALA.
FT   DOMAIN      721    727       POLY-PRO.
FT   DISULFID    257    311       BY SIMILARITY.
FT   DISULFID    345    359       BY SIMILARITY.
FT   DISULFID    353    370       BY SIMILARITY.
FT   DISULFID    372    381       BY SIMILARITY.
FT   CARBOHYD     52     52       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     53     53       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    278    278       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    346    346       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC      1    233       MROVCCSALPPPPLEKGRCSSYSDSSSSSSSERSSSSSSSSS
FT                                ESGSSSRSSSNNSSISRPAPPEPRPQQOPQPRSPAARRAA
FT                                ARSRAAAAGGMRRDPAPGFSMLLFGVSLACYSPSLKSVODQ
FT                                AYKAPVVVEGRVQGLVPAGGSSSSNSTREPPASGRVALVKVL
FT                                DKWPLRSGGLQREQVISVGSCVPLERNQRYIFFLEPTEQPL
FT                                VFKTAFAPLDTNGKNLKKEVGKILCTDC -> MSESRRGR
FT                                GRGKHPEGRKREREPDPGEK (IN ISOFORM DON-
FT                                1R).
FT   VARSPLIC      1    241       MISSING (IN ISOFORM DON-1B).
FT   VARSPLIC    374    396       NGFFGQRCLEKLPLRIYMPDPKQ -> VGGTGDRCQOFAMV
FT   VARSPLIC    374    397       NGFFGQRCLEKLPLRIYMPDPKQK ->
FT                                NFS (IN ISOFORM 2)
FT   VARSPLIC    397    397       NGFFGQRCLEKLPLRIYMPDPKQK -> VGYTGDRCQQFAM
FT                                VNFSKHLGFELKE (IN ISOFORM 4).
FT                                K -> KHLGFELKE (IN ISOFORM 3 AND IN
FT                                ISOFORM 1B).
FT   VARSPLIC    397    422       KAELIYQKRVLITIGICVALLVVGIV ->
FT                                SSSOWSTSPSTLDLN (IN ISOFORM 6).
FT   VARSPLIC    423    850       MISSING (IN ISOFORM 6).
FT   VARSPLIC    397    426       KAELIYQKRVLITIGICVALLVVGIVCVVA -> SVLWDTP
FT                                GTGVSSSOWSTSPKRSCRRRGS (IN ISOFORM 5).
FT   VARSPLIC    427    850       MISSING (IN ISOFORM 5).
SQ   SEQUENCE   850 AA;  91678 MW;  7124C0894435FD2F4 CRC64;
```

Query Match 32.5%; Score 90; DB 1; Length 850;

Best Local Similarity 34.9%;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

```
Qy     1 HFKPCRDKDLAYCLNDGECFYIETLLTGSHKHCRCKEGYQGVRC  43
           |  |:|  ||:||| |::| | |:|    |  |:| | || | ||
Db   341 HARKCNETAKSYCVNGGVCYYIEGI--NQLSCKCPNGFFGQRC  381
```

RESULT 11
NRG2_RAT

```
ID   NRG2_RAT     STANDARD;      PRT;   868 AA.
AC   035569; 035570; 035571; 035572; 035073;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pro-neuregulin-2, membrane-bound isoform precursor (Pro-NRG2) (Contains: Neuregulin-2 (NRG-2)
DE   (Neural- and thymus-derived activator for ERBB kinases) (NTAK)].
GN   NRG2 OR NTAK.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 128-162, AND ALTERNATIVE SPLICING.
RX   MEDLINE=98006324; PubMed=9348101;
RA   Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,
RA   Nakagawa T., Miyagawa J., Matsushita N., Nagatsu T., Taniguchi N.,
RA   Ishiguro H.;
RT   "A novel brain-derived member of the epidermal growth factor family
RT   that interacts with ErbB3 and ErbB4.";
RL   J. Biochem. 122:675-680(1997).
RN   [2]
RP   SEQUENCE OF 109-868 FROM N.A. (ISOFORMS NRG2-ALPHA AND NRG2-BETA).
RC   TISSUE=Cerebellum;
RX   MEDLINE=97311397; PubMed=9168114;
RA   Chang H., Riese D.J. II, Gilbert W., Stern D.F., McMahan U.J.;
RT   "Ligands for ErbB-family receptors encoded by a neuregulin-like
RT   gene.";
RL   Nature 387:509-512(1997).
CC   -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC       RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC       RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC       ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
CC       HETERODIMERIZATION WITH THE EGF RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC       A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC       MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 7 ISOFORMS; NTAK-ALPHA1 (SHOWN
CC       HERE), NTAK-ALPHA2A, NTAK-ALPHA2B/NTAK-ALPHA2-1P, NTAK-BETA, NTAK-
CC       GAMMA, NRG2-ALPHA AND NRG2-BETA, ARE PRODUCED BY ALTERNATIVE
CC       SPLICING. THE ALPHA-TYPE AND BETA-TYPE DIFFER IN THE EGF-LIKE
CC       DOMAIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOST PARTS OF THE BRAIN.
CC       ESPECIALLY THE OLFACTORY BULB AND CEREBELLUM WHERE IT LOCALIZES IN
CC       GRANULE AND PURKINJE CELLS. IN THE HIPPOCAMPUS, FOUND IN THE
CC       GRANULE CELLS OF THE DENTRATE GYRUS. IN THE BASAL FOREBRAIN, FOUND
CC       IN THE CHOLINERGIC CELLS. IN THE EMBRYO, WEAKLY DETECTABLE IN
CC       THE MOTOR TRIGEMINAL NUCLEUS. NOT DETECTED IN THE HYPOTHALAMUS.
CC       ALSO FOUND IN THE LIVER AND IN THE THYMUS. NOT DETECTED IN HEART,
CC       ADRENAL GLAND, OR TESTIS.
CC   -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, EXPRESSED IN THE BRAIN OF
CC       E11.5 EMBRYOS WHERE IT IS FOUND IN THE TELENCEPHALON, BUT NOT IN
CC       THE HINDBRAIN. NOT FOUND IN THE HEART. IN THE ADULT, FOUND IN
CC       BRAIN AND THYMUS.
CC   -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC       OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC       PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC       DIMERIZATION (BY SIMILARITY).
CC   -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC       DOMAIN (BY SIMILARITY).
CC   -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC       EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC       FORM (BY SIMILARITY).
```

Qy   1 HFKPCRDKDLAYCLNDGECFYIETLLTGSHKHCRCKEGYQGVRC 43

Db   341 HARKCNETAKSYCVNGGVCYYIEGI--NQLSCKCPNGFFGQRC 381

Db   341 HARKCNETAKSYCVNGGVCYYIEGI--NQLSCKCPNGFFGQRC 381

```
CC   -!- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement.(See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; D89995; BAA23344.1; -.
DR   EMBL; D89996; BAA23345.1; -.
DR   EMBL; D89997; BAA23346.1; -.
DR   EMBL; D89998; BAA23347.1; -.
DR   EMBL; D89999; BAA23348.1; -.
DR   HSSP; Q12784; 1HRE.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_C2.
DR   InterPro; IPR002154; Neuregulin.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; Ig; 1.
DR   Pfam; PF02158; Neuregulin; 2.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW   Transmembrane; Multigene family; Alternative splicing.
FT   PROPEP        1    127
FT   CHAIN       128    868
FT   CHAIN       128    428
FT   DOMAIN      128    429
FT   TRANSMEM    430    450
FT   DOMAIN      451    868       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      266    334       EGF-LIKE.
FT   DOMAIN      346    356       SER/THR-RICH.
FT   DOMAIN      357    398       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN       22     32       EGF-LIKE.
FT   DOMAIN       35     45       POLY-SER.
FT   DOMAIN       32     32       POLY-SER.
FT   DOMAIN       56     59       POLY-THR.
FT   DOMAIN      103    106       POLY-ALA.
FT   DOMAIN      739    745       POLY-PRO.
FT   DISULFID    273    327       BY SIMILARITY.
FT   DISULFID    361    375       BY SIMILARITY.
FT   DISULFID    369    386       BY SIMILARITY.
FT   DISULFID    388    397       BY SIMILARITY.
FT   CARBOHYD     33     33       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     34     34       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    163    163       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    362    362       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC      1    108       MISSING (IN ISOFORM NRG2-BETA).
FT   VARSPLIC    220    222       PLV -> FFF (IN ISOFORM NTAK-ALPHA2-1P).
FT   VARSPLIC    388    388       C -> G (IN ISOFORM NTAK-GAMMA).
FT   VARSPLIC    389    868       MISSING (IN ISOFORM NTAK-GAMMA).
FT   VARSPLIC    390    412       NGFFGQRCLEKLPLRLYMPDPKQ -> VGYTGDRCQQFAMV
FT                                NFS (IN ISOFORM NRG2-BETA).
FT   VARSPLIC    390    421       NGFFGQRCLEKLPLRLYMPDPKRQKHLGFELKE -> VGYTG
FT                                DRCQQFAMVNFSK (IN ISOFORM NTAK-BETA).
FT   VARSPLIC    414    414       MISSING (IN ISOFORM NTAK-ALPHA2A).
FT   VARSPLIC    414    439       MISSING (IN ISOFORM NTAK-ALPHA2B).
FT   VARSPLIC    440    868       HLGFELKEAEELYQKRVLTITGICVA -> SVLMDTPGTGV
FT                                SSSOWSTSPSTLDLN (IN ISOFORM NRG2-ALPHA).
FT   CONFLICT    117    117       MISSING (IN ISOFORM NRG2-ALPHA).
FT   CONFLICT    724    724       S -> F (IN REF. 2).
FT                                R -> H (IN REF. 2).
SO   SEQUENCE   868 AA;  93776 MW;  3C7D4D94DBE64DE2 CRC64;
```

```
                                                         RESULT 12
                                                         EREG_HUMAN

Qy           1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRC  43
                   :|| : :||:| ||  |: : ||:| |  ||| |   |    |
Db         357  HARKCNETAKSYCVNGGVCYYIEGI--NQLSCKCPNGFFGQRC  397

ID   EREG_HUMAN             STANDARD;      PRT;   169 AA.
AC   O14944;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Epiregulin precursor.
GN   EREG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Colorectal adenocarcinoma;
RX   MEDLINE=97479200; PubMed=9337852;
RA   Toyoda H., Komurasaki T., Uchida D., Morimoto S.;
RT   "Distribution of mRNA for human epiregulin, a differentially expressed
RT   member of the epidermal growth factor family.";
RL   Biochem. J. 326:69-75(1997).
CC   -!- FUNCTION: MAY BE A MEDIATOR OF LOCALIZED CELL PROLIFERATION. AS A
CC       MITOGEN IT MAY STIMULATE CELL PROLIFERATION AND/OR ANGIOGENESIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);
CC       EXTRACELLULAR (MATURE FORM).
CC   -!- TISSUE SPECIFICITY: IN NORMAL ADULTS EXPRESSED PREDOMINANTLY IN THE
CC       PLACENTA AND PERIPHERAL BLOOD LEUCOYTES. HIGH LEVELS WERE DETECTED
CC       IN CARCINOMAS OF THE BLADDER, LUNG, KIDNEY AND COLON.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement.(See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; D30783; BAA22146.1; -.
DR   MIM; 602061; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001336; EGF_1.
DR   PRINTS; PR00009; EGFTGF.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; Mitogen; Glycoprotein; EGF-like domain; Transmembrane;
KW   Signal.
FT   SIGNAL        1     29       POTENTIAL.
FT   PROPEP       30     59       POTENTIAL.
FT   CHAIN        60    108       EPIREGULIN.
FT   PROPEP      109    169       REMOVED IN MATURE FORM (BY SIMILARITY).
FT   DOMAIN       60    119       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    120    140       POTENTIAL.
FT   DOMAIN      141    169       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      145    152       ARG/LYS-RICH (BASIC).
FT   CARBOHYD     47     47       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DOMAIN       64    104       EGF-LIKE.
FT   DISULFID     68     81       BY SIMILARITY.
FT   DISULFID     76     92       BY SIMILARITY.
FT   DISULFID     94    103       BY SIMILARITY.
SO   SEQUENCE   169 AA;  19044 MW;  17F3926ADFB2BDEE CRC64;
```

Query Match   32.5%; Score 90; DB 1; Length 868;
Best Local Similarity 34.9%; Pred. No. 0.002;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRC 43
                   :|| : :||:| || |: : ||:| | ||| | | |
Db   357 HARKCNETAKSYCVNGGVCYYIEGI--NQLSCKCPNGFFGQRC 397

Query Match   32.3%; Score 89.5; DB 1; Length 169;
Best Local Similarity 40.5%; Pred. No. 0.0005;

Matches 17; Conservative 7; Mismatches 15; Indels 3; Gaps 2;

RESULT 13
GRFA_SFVKA

```
Qy    5 CRDKDLAYCLNDGECFVIETLTGGSHKHCRCKEGYQGVRCDQF 46
          |||:| |:| :| ||| |||| |||||| ||| ||| |
Db   68 CSSDMNGYCLH-GQCIYLVDM--SQNYCRCEVGYTGVRCEHF 106
```

ID   GRFA_SFVKA   STANDARD;   PRT;   80 AA.
AC   P08441;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Growth factor.
OS   Shope fibroma virus (strain Kasza) (SFV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Leporipoxvirus.
OX   NCBI_TaxID=10272;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87172751; PubMed=3031480;
RA   Chang W., Upton C., Hu S.-L., Purchio A.F., McFadden G.;
RT   "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains
RT   a growth factor gene with sequence similarity to those encoding
RT   epidermal growth factor and transforming growth factor alpha.";
RL   Mol. Cell. Biol. 7:535-540(1987).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; M15921; AAA46873.1; -.
DR   PIR; A26723; EGV2SF.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001336; EGF_1.
DR   PRINTS; PR00009; EGFTGF.
DR   SMART; SM00181; EGF; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   EGF-like domain; Growth factor; Glycoprotein.
FT   DOMAIN   29   73
FT   DISULFID   33   47   BY SIMILARITY.
FT   DISULFID   41   61   BY SIMILARITY.
FT   DISULFID   63   72   BY SIMILARITY.
FT   CARBOHYD   44   44   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   54   54   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   80 AA;   9210 MW;   C4BD3OE878D2ED58 CRC64;

RESULT 14
NTC1_MOUSE

```
Qy    1 HFKPCRDLAYCLNDGECFVIE----TLTGSHKHCRCKEGYQGVRCDQFL 47
         |:||:||||||||||||||||   :| || |||||:||||:| |||:|:
Db   29 HVKVCNHDYENYCLNNGTCFTIALDNVSIT---PFCVCRINYEGSRC-QFI 75
```

Query Match   30.7%;   Score 85;   DB 1;   Length 80;
Best Local Similarity   39.2%;   Pred. No. 0.00087;
Matches 20;   Conservative 6;   Mismatches 17;   Indels 8;   Gaps 3;

ID   NTC1_MOUSE   STANDARD;   PRT;   2531 AA.
AC   001705;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neurogenic locus notch homolog protein 1 precursor (MOTCH protein).
GN   NOTCH1 OR MOTCH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=93194170; PubMed=8449489;
RA   Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RA   Copeland N.G., Gridley T.;
RT   "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT   homolog of Drosophila Notch.";
RL   Genomics 15:259-264(1993).
RN   [2]
RP   SEQUENCE OF 1551-2170 FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=93048835; PubMed=1425352;
RA   Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RA   Greenspan R.J., McMahon A.P., Gridley T.;
RT   "Expression pattern of Motch, a mouse homolog of Drosophila Notch,
RT   suggests an important role in early postimplantation mouse
RT   development.";
RL   Development 115:737-744(1992).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC   -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC   -!- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC   -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; Z11886; CAA77941.1; -.
DR   HSSP; P00740; 1EDM.
MGD; MGI:97363; Notch1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR000800; Notch.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00008; EGF; 35.
DR   Pfam; PF00066; notch; 3.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR01452; NOTCH.
DR   PROSITE; PS01187; EGF_CA; 21.
DR   SMART; SM00248; ANK; 3.
DR   SMART; SM00179; EGF_CA; 23.
DR   SMART; SM00004; NL; 2.
DR   SMART; SM00001; EGF_like; 11.
DR   SMART; SM00008; EGF; 35.
DR   PROSITE; PS50088; ANK_REPEAT; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 22.
DR   PROSITE; PS00022; EGF_1; 34.
DR   PROSITE; PS01186; EGF_2; 27.
DR   PROSITE; PS01187; EGF_CA; 21.
KW   Transmembrane; Signal; Glycoprotein.
KW   Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
FT   SIGNAL   1   18
FT   CHAIN   19   2531   NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT   DOMAIN   19   1725   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1726   1746   POTENTIAL.
FT   DOMAIN   1747   2531   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   20   58   EGF-LIKE 1.
FT   DOMAIN   59   99   EGF-LIKE 2.
FT   DOMAIN   102   139   EGF-LIKE 3.
FT   DOMAIN   140   176   EGF-LIKE 4.

```
FT   DOMAIN       178    216       EGF-LIKE 5.
FT   DOMAIN       218    255       EGF-LIKE 6.
FT   DOMAIN       257    293       EGF-LIKE 7.
FT   DOMAIN       295    333       EGF-LIKE 8.
FT   DOMAIN       335    371       EGF-LIKE 9.
FT   DOMAIN       372    410       EGF-LIKE 10.
FT   DOMAIN       412    450       EGF-LIKE 11.
FT   DOMAIN       452    488       EGF-LIKE 12.
FT   DOMAIN       490    526       EGF-LIKE 13.
FT   DOMAIN       528    564       EGF-LIKE 14.
FT   DOMAIN       566    601       EGF-LIKE 15.
FT   DOMAIN       603    639       EGF-LIKE 16.
FT   DOMAIN       641    676       EGF-LIKE 17.
FT   DOMAIN       678    714       EGF-LIKE 18.
FT   DOMAIN       716    751       EGF-LIKE 19.
FT   DOMAIN       753    789       EGF-LIKE 20.
FT   DOMAIN       791    827       EGF-LIKE 21.
FT   DOMAIN       829    867       EGF-LIKE 22.
FT   DOMAIN       869    905       EGF-LIKE 23.
FT   DOMAIN       907    943       EGF-LIKE 24.
FT   DOMAIN       945    981       EGF-LIKE 25.
FT   DOMAIN       983   1019       EGF-LIKE 26.
FT   DOMAIN      1021   1057       EGF-LIKE 27.
FT   DOMAIN      1059   1095       EGF-LIKE 28.
FT   DOMAIN      1097   1143       EGF-LIKE 29.
FT   DOMAIN      1145   1181       EGF-LIKE 30.
FT   DOMAIN      1183   1219       EGF-LIKE 31.
FT   DOMAIN      1221   1265       EGF-LIKE 32.
FT   DOMAIN      1267   1305       EGF-LIKE 33.
FT   DOMAIN      1307   1346       EGF-LIKE 34.
FT   DOMAIN      1348   1384       EGF-LIKE 35.
FT   DOMAIN      1387   1426       EGF-LIKE 36.
FT   DOMAIN      1449   1462       CYS-RICH.
FT   REPEAT      1481   1522       LIN/NOTCH 1.
FT   REPEAT      1523   1562       LIN/NOTCH 2.
FT   REPEAT      1917   1947       LIN/NOTCH 3.
FT   REPEAT      1949   1979       ANK 1.
FT   REPEAT      1983   2012       ANK 2.
FT   REPEAT      2016   2045       ANK 3.
FT   REPEAT      2049   2078       ANK 4.
FT   DOMAIN        24     37       ANK 5.

FT   CA_BIND      216           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      255           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      293           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      333           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      371           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      410           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      450           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      488           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      526           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      564           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      601           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      639           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      676           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      714           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      751           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      789           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      827           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      867           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      905           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      943           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND      981           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND     1019           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND     1057           CALCIUM-BINDING (POTENTIAL).
FT   CA_BIND     1095           CALCIUM-BINDING (POTENTIAL).

FT   DISULFID      31     46       BY SIMILARITY.
FT   DISULFID      63     74       BY SIMILARITY.
FT   DISULFID      68     87       BY SIMILARITY.
FT   DISULFID      89     98       BY SIMILARITY.
FT   DISULFID     106    117       BY SIMILARITY.
FT   DISULFID     111    127       BY SIMILARITY.
FT   DISULFID     129    138       BY SIMILARITY.
FT   DISULFID     144    155       BY SIMILARITY.
FT   DISULFID     149    164       BY SIMILARITY.
FT   DISULFID     166    175       BY SIMILARITY.
FT   DISULFID     182    195       BY SIMILARITY.
FT   DISULFID     189    204       BY SIMILARITY.
FT   DISULFID     206    215       BY SIMILARITY.
FT   DISULFID     222    233       BY SIMILARITY.
FT   DISULFID     227    243       BY SIMILARITY.
FT   DISULFID     245    254       BY SIMILARITY.
FT   DISULFID     261    272       BY SIMILARITY.
FT   DISULFID     266    281       BY SIMILARITY.
FT   DISULFID     283    292       BY SIMILARITY.
FT   DISULFID     299    312       BY SIMILARITY.
FT   DISULFID     306    321       BY SIMILARITY.
FT   DISULFID     323    332       BY SIMILARITY.
FT   DISULFID     339    350       BY SIMILARITY.
FT   DISULFID     344    359       BY SIMILARITY.
FT   DISULFID     361    370       BY SIMILARITY.
FT   DISULFID     376    387       BY SIMILARITY.
FT   DISULFID     381    398       BY SIMILARITY.
FT   DISULFID     400    409       BY SIMILARITY.
FT   DISULFID     416    429       BY SIMILARITY.
FT   DISULFID     423    438       BY SIMILARITY.
FT   DISULFID     440    449       BY SIMILARITY.
FT   DISULFID     456    467       BY SIMILARITY.
FT   DISULFID     461    476       BY SIMILARITY.
FT   DISULFID     478    487       BY SIMILARITY.
FT   DISULFID     494    505       BY SIMILARITY.
FT   DISULFID     499    514       BY SIMILARITY.
FT   DISULFID     516    525       BY SIMILARITY.
FT   DISULFID     532    543       BY SIMILARITY.
FT   DISULFID     537    552       BY SIMILARITY.
FT   DISULFID     554    563       BY SIMILARITY.
FT   DISULFID     570    580       BY SIMILARITY.
FT   DISULFID     575    589       BY SIMILARITY.
FT   DISULFID     591    600       BY SIMILARITY.
FT   DISULFID     607    618       BY SIMILARITY.
FT   DISULFID     612    627       BY SIMILARITY.
FT   DISULFID     629    638       BY SIMILARITY.
FT   DISULFID     645    655       BY SIMILARITY.
FT   DISULFID     650    664       BY SIMILARITY.
FT   DISULFID     666    675       BY SIMILARITY.
FT   DISULFID     682    693       BY SIMILARITY.
FT   DISULFID     687    702       BY SIMILARITY.
FT   DISULFID     704    713       BY SIMILARITY.
FT   DISULFID     720    730       BY SIMILARITY.
FT   DISULFID     725    739       BY SIMILARITY.
FT   DISULFID     741    750       BY SIMILARITY.
FT   DISULFID     757    768       BY SIMILARITY.
FT   DISULFID     762    777       BY SIMILARITY.
FT   DISULFID     779    788       BY SIMILARITY.
FT   DISULFID     795    806       BY SIMILARITY.
FT   DISULFID     800    815       BY SIMILARITY.
FT   DISULFID     817    826       BY SIMILARITY.
FT   DISULFID     833    844       BY SIMILARITY.
FT   DISULFID     838    855       BY SIMILARITY.
FT   DISULFID     857    866       BY SIMILARITY.
FT   DISULFID     873    884       BY SIMILARITY.
FT   DISULFID     878    893       BY SIMILARITY.
FT   DISULFID     895    904       BY SIMILARITY.
FT   DISULFID     911    922       BY SIMILARITY.
FT   DISULFID     916    931       BY SIMILARITY.
FT   DISULFID     933    942       BY SIMILARITY.
FT   DISULFID     987    998       BY SIMILARITY.
FT   DISULFID     992   1007       BY SIMILARITY.
FT   DISULFID    1009   1018       BY SIMILARITY.
FT   DISULFID    1025   1036       BY SIMILARITY.
FT   DISULFID    1030   1045       BY SIMILARITY.
FT   DISULFID    1047   1056       BY SIMILARITY.
FT   DISULFID    1063   1074       BY SIMILARITY.
```

RESULT 15
BTC_MOUSE

ID   BTC_MOUSE   STANDARD;   PRT;   177 AA.
AC   Q05928;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Betacellulin precursor (BTC).
GN   BTC OR BCN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 32-54; 64-71 AND 75-111.
RC   TISSUE=Pancreas;
RX   MEDLINE=93206093; PubMed=8456283;

```
Qy      9 DLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCD 44
          ||||||||  ||   |:|:|  | |:|:  |||:|:||:|||
Db   1064 DSAPCKNGGRCW----QTNTQYHCECRSGWTGVNCD 1095
```

Query Match   30.0%; Score 83; DB 1; Length 2531;
Best Local Similarity 41.7%; Pred. No. 0.037;
Matches   15; Conservative   4; Mismatches   13; Indels   4; Gaps   1;

```
RA   Shing Y., Christofori G., Hanahan D., Ono Y., Sasada R.,
RA   Igarashi K., Folkman J.;
RT   "Betacellulin: a mitogen from pancreatic beta cell tumors.";
RL   Science 259:1604-1607(1993).
CC   -!- FUNCTION: POTENT MITOGEN FOR RETINAL PIGMENT EPITHELIAL CELLS
CC       AND VASCULAR SMOOTH MUSCLE CELLS. THE EFFECTS OF BETACELLULIN
CC       ARE PROBABLY MEDIATED BY THE EGF RECEPTOR AND OTHER RELATED
CC       RECEPTORS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);
CC       EXTRACELLULAR (MATURE FORM).
CC   -!- TISSUE SPECIFICITY: FOUND IN SEVERAL MOUSE TISSUES INCLUDING
CC       KIDNEY, UTERUS AND LIVER AS WELL AS IN BETA TUMOR CELL LINE AND
CC       MCF-7 CELLS. IT IS NOT DETECTED IN THE BRAIN.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L08394; AAA40511.1; -.
DR   PIR; A37408; A37408.
DR   HSSP; P01135; 3TGF.
DR   MGD; MGI:99439; Btc.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001336; EGF_1.
DR   Pfam; PF00008; EGF; 1.
DR   PRINTS; PR00009; EGFTGF.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   Growth factor; Mitogen; Glycoprotein; EGF-like domain; Transmembrane;
KW   Signal.
FT   SIGNAL           1     31
FT   CHAIN           32    177       BETACELLULIN.
FT   PROPEP         112    177       REMOVED IN MATURE FORM.
FT   DOMAIN          32    118       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       119    139       POTENTIAL.
FT   DOMAIN         140    177       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN          65    105       EGF-LIKE.
FT   DOMAIN         146    153       ARG/LYS-RICH (BASIC).
FT   DISULFID        69     82       BY SIMILARITY.
FT   DISULFID        77     93       BY SIMILARITY.
FT   DISULFID        95    104       BY SIMILARITY.
FT   CARBOHYD        34     34       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD        42     42       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD        52     52       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE       177 AA;  19664 MW;  066BB34F0E13F82B CRC64;
```

Query Match   29.8%; Score 82.5; DB 1; Length 177;
Best Local Similarity   39.1%; Pred. No. 0.0036;
Matches   18; Conservative   9; Mismatches   14; Indels   5; Gaps   3;

```
Qy          1 HFKPCRDKDLAYCLNDGEC-FVIETLTGSHKHCRCKEGYOGVRCDQ 45
              || || |:   ||::|| ::||:|:|    :|  |::||::||
Db         65 HFSRCPKQYKHYCIH-GRCRFVVDEQTPS---CICEKGYFGARCER 106
```

OM protein - protein search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: November 7, 2002, 09:59:08 ; Search time 19 seconds
                 (without alignments)
                 237.695 Million cell updates/sec

Title: US-09-480-977-4
Perfect score: 277
Sequence: 1 HFKPCRDKLAYCLNDGECF.........SHKHCRCKEGYQGVRCDQFL 47

Scoring table: Gapop 10.0 , Gapext 0.5
                 BLOSUM62

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_71:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	277	100.0	713	2	T44447	neuregulin-3 [impo
2	113.5	41.0	125	2	I38405	neu differentiatio
3	113.5	41.0	462	2	I38404	neu differentiatio
4	113.5	41.0	640	2	A43273	heregulin precurso
5	111.5	40.3	639	2	I61719	neu differentiatio
6	105.5	38.1	125	2	S62676	heregulin isoform
7	104.5	37.7	175	2	I38408	heregulin isoform
8	104.5	37.7	241	2	S32359	neu differentiatio
9	104.5	37.7	241	2	S32359	glial growth facto
10	104.5	37.7	296	2	A56943	sensory/motor neur
11	104.5	37.7	422	2	S32357	glial growth facto
12	104.5	37.7	637	2	C43273	heregulin precurso
13	104.5	37.7	645	2	B43273	heregulin precurso
14	102.5	37.0	230	2	A56210	neu differentiatio
15	102.5	37.0	230	2	I61718	neu differentiatio
16	102.5	37.0	662	2	I61722	neu differentiatio
17	98.5	35.6	602	2	A45769	acetylcholine rece
18	92.5	33.4	2180	2	T29764	hypothetical prote
19	90	32.5	850	2	JC5700	ErbB kinase activa
20	90	32.5	860	2	JC5702	ErbB kinase activa
21	90	32.5	868	2	JC5701	ErbB kinase activa
22	85	30.7	80	2	EGVZSF	growth factor - ra
23	84.5	30.5	1220	2	A56136	jagged protein pre
24	83	30.0	46	2	JT0747	epiregulin - rat
25	83	30.0	162	2	S68401	epiregulin precurs
26	83	30.0	861	2	A48825	epiregulin precurs
27	83	30.0	2531	2	A46019	Notch homolog Motc
28	82.5	29.8	177	2	A37408	Notch-1 protein -
29	82	29.6	85	1	EGVZM1	betacellulin precu
30	82	29.6	230	2	A44074	probable EGF-like
31	82	29.6	264	2	T22380	hypothetical prote
32	80.5	29.1	178	2	JC1467	betacellulin precu
33	80	28.9	907	2	T27317	hypothetical prote
34	79.5	28.7	140	1	WMVZ9	growth factor - va
35	79.5	28.7	140	1	T30766	growth factor - va
36	79.5	28.7	142	1	WMVZ3C	growth factor - va
37	79.5	28.7	159	1	WFRT1	transforming growt
38	79.5	28.7	159	1	S27195	transforming growt
39	79	28.5	2531	2	S18108	notch protein homo
40	79	28.5	1207	1	EGHU	epidermal growth homo
41	78	28.2	722	1	I48324	growth factor - va
42	77.5	28.0	159	1	I57997	DELTA-like 1 - mou
43	77.5	28.0	2352	2	T30201	Notch homolog prot
44	76.5	27.6	1372	2	T25933	hypothetical prote
45	76	27.4	482	2	JC5092	E-selectin - pig

ALIGNMENTS

RESULT 1
T44447
neuregulin-3 [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T44447
R;Zhang, D.; Sliwkowski, M.X.; Mark, M.; Frantz, G.; Akita, R.; Sun, Y.; Hillan, K.
Proc. Natl. Acad. Sci. U.S.A. 94, 9562-9567, 1997
A;Title: Neuregulin-3 (NRG3): A novel neural tissue-enriched protein that binds and
A;Reference number: Z22773; MUID:97420720
A;Accession: T44447
A;Molecule type: mRNA
A;Residues: 1-713 <ZHA>
A;Cross-references: EMBL:AF010130; NID:g2429163; PIDN:AAB70914.1; PID:g2429164
A;Gene: NRG3

Query Match   100.0%; Score 277; DB 2; Length 713;
Best Local Similarity   100.0%;
Matches   47; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 HFKPCRDKLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 288 HFKPCRDKLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334

RESULT 2
I38405
neu differentiation factor - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C;Accession: I38405
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen,
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiati
A;Reference number: I38405
A;Accession: A56210; MUID:94158863
A;Molecule type: mRNA
A;Residues: 1-125 <RES>
A;Cross-references: EMBL:U02327; NID:g408404; PIDN:AAA19952.1; PID:g408405

Query Match   41.0%; Score 113.5; DB 2; Length 125;
Best Local Similarity   34.8%;
Matches   16; Conservative   14; Mismatches   15; Indels   1; Gaps   1;

Qy   1 HFKPCRDKLAYCLNDGECFVIETLTGSRKH-CRCKEGYQGVRCDQ 45
         | ::| :|:|||||::|:: |: ::|:|| ||:|:: |:|| :
Db   56 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 101

RESULT 3

I38404
neu differentiation factor - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 05-Nov-1999
C;Accession: I38404
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A;Reference number: A56210; MUID:94158863
A;Accession: I38404
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-462 <RES>
A;Cross-references: EMBL:U02326; NID:g408402; PIDN:AAA19951.1; PID:g408403

```
Query Match           41.0%;  Score 113.5;  DB 2;  Length 462;
Best Local Similarity 34.8%;  Pred. No. 1.6e-05;
Matches 16;  Conservative 14;  Mismatches 15;  Indels 1;  Gaps 1;

Qy    1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
           |:|: |||||| ||:||| ||||::||:|| |||:|: |||||||
Db  178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 223
```

RESULT 4

A43273
heregulin precursor, splice form alpha - human
N;Alternate names: breast cancer cell differentiation factor p45; Neu differentiation fa
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 11-Jan-2000
C;Accession: A43273; A48498; A38155
R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansur
Science 256, 1205-1210, 1992
A;Title: Identification of heregulin, a specific activator of p185(erbB2).
A;Reference number: A43273; MUID:92271253
A;Accession: A43273
A;Molecule type: protein
A;Residues: 20-21,'X',23-24,'XX',27-28 <CUL>
R;Peles, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Yarde
Cell 69, 205-216, 1992
A;Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induce
A;Reference number: A48498; MUID:93367731
A;Accession: A48498
A;Molecule type: protein
A;Residues: 'X',15-16,'X',18-20,'RG',23-24,'GP',27,'E',29,'XP',32-36 <PEL>
R;Culouscou, J.M.; Plowman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.
J. Biol. Chem. 268, 18407-18410, 1993
A;Title: Characterization of a breast cancer cell differentiation factor that specifica
A;Reference number: A38155; MUID:92208945
A;Accession: A38155
A;Molecule type: protein
A;Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26
A;Note: sequence extracted from NCBI backbone (NCBIP:103250)
A;Molecule type: nucleic acid sequence not shown; not compared with conceptual translation
A;Residues: 1-640 <HOL>
A;Cross-references: GDB:132656; OMIM:142445
A;Gene: GDB:HGL
A;Genetics:
A;Map position: 8p22-8p11
A;Superfamily: unassigned EGF-related proteins; EGF homology
F;182-221/Domain: EGF homology <EGF>

```
Query Match           41.0%;  Score 113.5;  DB 2;  Length 640;
Best Local Similarity 34.8%;  Pred. No. 2.1e-05;
Matches 16;  Conservative 14;  Mismatches 15;  Indels 1;  Gaps 1;

Qy    1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
Db  178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 223
```

RESULT 5

I61719
neu differentiation factor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 01-Dec-2000
C;Accession: I61719; I61723; I61716; I61717; I61724; A38220
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiatio
A;Reference number: A56210; MUID:94158863
A;Accession: I61719
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-639 <RE2>
A;Cross-references: EMBL:U02323; NID:g408396; PIDN:AAA19948.1; PID:g408397
A;Accession: I61716
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-422 <RE5>
A;Cross-references: EMBL:U02316; NID:g408382; PIDN:AAA19941.1; PID:g408383
A;Accession: I61717
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-422, 'H', 'NL', '637-638,' ELRRNKAYRSKCMQIOLSATHLRPSSITHLGFIL', <RE4>
A;Cross-references: EMBL:U02317; NID:g408384; PIDN:AAA19942.1; PID:g408385
A;Accession: I61723
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-639 <RES>
A;Cross-references: EMBL:U02319; NID:g408388; PIDN:AAA19944.1; PID:g408389
A;Accession: I61724
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-422 <RE3>
A;Cross-references: EMBL:U02324; NID:g408398; PIDN:AAA19949.1; PID:g408399
A;Molecule type: mRNA
A;Residues: 1-422 <RE5>
A;Accession: A38220
R;Peles, E.; Bacus, R.; Cupples, R.; Suggs, S.V.; Bacus, S.S.; Luo, Y.; Trail, G.; Hu,
Cell 69, 559-572, 1992
A;Title: Neu differentiation factor: a transmembrane glycoprotein containing an EGF
A;Reference number: A38220; MUID:92257596
A;Accession: A38220
A;Note: sequence extracted from NCBI backbone (NCBIN:101767, NCBIP:101768)
A;Superfamily: unassigned EGF-related proteins; EGF homology

```
Query Match           40.3%;  Score 111.5;  DB 2;  Length 639;
Best Local Similarity 34.8%;  Pred. No. 3.5e-05;
Matches 16;  Conservative 13;  Mismatches 16;  Indels 1;  Gaps 1;

Qy    1 HFKPCRDKLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
           |:|: |||||| ||:||| ||||::|| || |||:|: |||||||
Db  178 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCQPGFTGARCTE 223
```

RESULT 6

S62676
heregulin isoform alpha 2 - human (fragments)
N;Alternate names: differentiation factor neu isoform alpha 2
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C;Accession: S62676
R;Hara, S.; Liu, N.; Meng, S.Y.; Lu, H.S.
Biochim. Biophys. Acta 1292, 168-176, 1996
A;Title: Isolation and structural characterization of recombinant human neu differe
A;Reference number: S62676; MUID:96139341
A;Accession: S62676
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-6;7-16;17-30;31-38;39-58;59-92;93-120;121-125 <HAR>
A;Superfamily: proto-oncogene
C;Keywords: proto-oncogene

```
Qy          Query Match         38.1%;    Score 105.5;   DB 2;   Length 125;
            Best Local Similarity    32.6%;    Pred. No. 4.2e-05;
            Matches    15;    Conservative    14;    Mismatches    16;    Indels    1;    Gaps    1;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
            | :|: .|||.|||||:||: ||| |:|||   :|| | ||:||
Db   75 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPQFTGARCTE 120

Db   .75 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPQFTGARCTE 120

RESULT 7
138408
neu differentiation factor - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998
C;Accession: 138408
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A;Reference number: A56210; MUID: MUID:94158863
A;Accession: 138408
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-175 <RES>
A;Cross-references: EMBL:U02330; NID:g408410; PIDN:AAA19955.1; PID:g408411
C;Superfamily: unassigned EGF-related proteins; EGF homology
F;116-155/Domain: EGF homology <EGF>

            Query Match         37.8%;    Score 104.5;   DB 2;   Length 175;
            Best Local Similarity    31.2%;    Pred. No. 7.3e-05;
            Matches    15;    Conservative    14;    Mismatches    18;    Indels    1;    Gaps    1;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|: .|||.|||||:||: ||| |:|||   :|| | ||:|||
Db  112 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPNEFTGDRCQNYV 159

RESULT 8
S32359
glial growth factor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997    #sequence_revision 01-Aug-1997    #text_change 11-Jan-2000
C;Accession: S32359
R;Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Hen
les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M.
Nature 362, 312-318, 1993
A;Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the n
A;Reference number: S32357; MUID:93205115
A;Accession: S32359
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-241 <MAR>
A;Cross-references: GB:L12259; NID:g289413; PIDN:AAA30540.1; PID:g289414
C;Superfamily: unassigned EGF-related proteins; EGF homology
F;182-221/Domain: EGF homology <EGF>

            Query Match         37.8%;    Score 104.5;   DB 2;   Length 241;
            Best Local Similarity    31.2%;    Pred. No. 9.6e-05;
            Matches    15;    Conservative    14;    Mismatches    18;    Indels    1;    Gaps    1;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|: .|||.|||||:||: ||| |:|||   :|| | ||:|||
Db  178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPNEFTGDRCQNYV 225

RESULT 9
D43273
heregulin precursor, splice form beta-3 - human
N;Alternate names: glial growth factor HRG-beta-3; neuregulin
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993    #sequence_revision 31-Dec-1993    #text_change 11-Jan-2000
C;Accession: D43273, S32358
R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansur
Science 256, 1205-1210, 1992
```

```
A;Title: Identification of heregulin, a specific activator of p185(erbB2).
A;Reference number: A43273; MUID:9221253
A;Accession: D43273
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptua
A;Molecule type: mRNA
A;Residues: 1-241 <HOL>
C;Genetics:
A;Gene: GDB:HGL; GGF
A;Cross-references: GDB:132656; OMIM:142445
A;Map position: 8p22-8p11
C;Superfamily: unassigned EGF-related proteins; EGF homology
F;182-221/Domain: EGF homology <EGF>

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|: .|||.|||||:||: ||| |:|||   :|| | ||:|||
Db  178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPNEFTGDRCQNYV 225

RESULT 10
A56943
sensory/motor neuron-derived factor - human
C;Species: Homo sapiens (man)
C;Date: 18-Aug-1995    #sequence_revision 18-Aug-1995    #text_change 11-Jan-2000
C;Accession: A56943
R;Ho, W.H.; Armanini, M.P.; Nuijens, A.; Phillips, H.S.; Osheroff, P.L.
J. Biol. Chem. 270, 14523-14532, 1995
A;Title: Sensory and motor neuron-derived factor. A novel heregulin variant highly
A;Reference number: A56943; MUID:95301541
A;Accession: A56943
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-296 <HOA>
A;Cross-references: GB:L41827; NID:g862422; PIDN:AAC41764.1; PID:g862423
C;Superfamily: unassigned EGF-related proteins; EGF homology
F;237-276/Domain: EGF homology <EGF>

            Query Match         37.8%;    Score 104.5;   DB 2;   Length 296;
            Best Local Similarity    31.2%;    Pred. No. 0.00011;
            Matches    15;    Conservative    14;    Mismatches    18;    Indels    1;    Gaps    1;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|: .|||.|||||:||: ||| |:|||   :|| | ||:|||
Db  233 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPNEFTGDRCQNYV 280

RESULT 11
S32357
glial growth factor - human
C;Species: Homo sapiens (man)
C;Date: 02-Dec-1993    #sequence_revision 10-Nov-1995    #text_change 11-Jan-2000
C;Accession: S32357
R;Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.
les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfiel
Nature 362, 312-318, 1993
A;Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in
A;Reference number: S32357; MUID:93205115
A;Accession: S32357
A;Status: preliminary
A;Molecule type: mRNA
```

A;Residues: 1-422 <MAT>
A;Cross-references: GB:L12260; NID:g292047; PIDN:AAB59622.1; PID:g292048
C;Superfamily: unassigned EGF-related proteins; EGF homology
F:363-402/Domain: EGF homology <EGF>

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|| :| ||||:|| ||| | ::| || : |:| ||| :| ||
Db   359 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCPNEFTGDRCQNYV 406

Query Match
Best Local Similarity   31.7%;   Score 104.5; DB 2; Length 422;
Matches   15; Conservative 14; Mismatches 18; Indels   1; Gaps   1;

RESULT 12
C43273
heregulin precursor, splice form beta-2 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Nov-2000
C;Accession: C43273; 138407
R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A;Title: Identification of heregulin, a specific activator of p185(erbB2).
A;Reference number: A43273; MUID:92271253
A;Accession: 138407
A;Molecule type: mRNA
A;Residues: 1-637 <HOL>
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 119-406 <RES>
A;Cross-references: EMBL:U02329; NID:g408408; PIDN:AAA19954.1; PID:g408409
C;Superfamily: unassigned EGF-related proteins; EGF homology
F:182-221/Domain: EGF homology <EGF>

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|| :| ||||:|| ||| | ::| || : |:| ||| :| ||
Db   178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCPNEFTGDRCQNYV 225

Query Match
Best Local Similarity   31.7%;   Score 104.5; DB 2; Length 637;
Matches   15; Conservative 14; Mismatches 18; Indels   1; Gaps   1;

RESULT 13
B43273
heregulin, splice form beta 1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Nov-2000
C;Accession: B43273; 138406
R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A;Title: Identification of heregulin, a specific activator of p185(erbB2).
A;Reference number: A43273; MUID:92271253
A;Accession: B43273
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-645 <HOL>
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A;Reference number: A56210; MUID:9415863
A;Accession: 138406

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-230 <RES>
A;Cross-references: EMBL:U02315; NID:g408380; PIDN:AAA19940.1; PID:g408381
C;Superfamily: unassigned EGF-related proteins; EGF homology

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|| :| ||||:|| ||| | ::| || : |:| ||| :| ||
Db   167 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 214

Query Match
Best Local Similarity   37.0%;   Score 102.5; DB 2; Length 230;
Matches   15; Conservative 13; Mismatches 19; Indels   1; Gaps   1;

RESULT 15
I61718
neu differentiation factor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 17-Nov-2000
C;Accession: I61718; I61721; I61720
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen,
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiatio
A;Status: Preliminary; translated from GB/EMBL/DDBJ
A;Accession: I61718
A;Molecule type: mRNA
A;Residues: 1-636 <RES>
A;Cross-references: EMBL:U02318; NID:g408386; PIDN:AAA19943.1; PID:g408387
A;Status: Preliminary; translated from GB/EMBL/DDBJ
A;Accession: I61721
A;Molecule type: mRNA
A;Residues: 1-444, 'A', 446-636 <RES>
A;Cross-references: EMBL:U02321; NID:g408392; PIDN:AAA19946.1; PID:g408393
A;Accession: I61720
A;Molecule type: mRNA
A;Residues: 1-298,386,'v',388,'TR',391 <RES>
A;Cross-references: EMBL:U02320; NID:g408390; PIDN:AAA19945.1; PID:g408391
C;Superfamily: unassigned EGF-related proteins; EGF homology

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'A', 95-418, 'F', 420-645 <RES>
A;Cross-references: EMBL:U02328; NID:g408406; PIDN:AAA19953.1; PID:g408407
C;Gene: GDB:HGL
C;Genetics:
A;Cross-references: GDB:132656; OMIM:142445
A;Map position: 8p22-8p11

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
            | :|| :| ||||:|| ||| | ::| || : |:| ||| :| ||
Db   178 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCPNEFTGDRCQNYV 225

Query Match
Best Local Similarity   37.7%;   Score 104.5; DB 2; Length 645;
Matches   15; Conservative 14; Mismatches 18; Indels   1; Gaps   1;

RESULT 14
A56210
neu differentiation factor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 26-May-2000
C;Accession: A56210
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, F
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiatio
A;Reference number: A56210; MUID:9415863
A;Accession: A56210
A;Molecule type: mRNA
A;Residues: 1-230 <RES>
A;Cross-references: EMBL:U02315; NID:g408380; PIDN:AAA19940.1; PID:g408381
C;Superfamily: unassigned EGF-related proteins; EGF homology

F:182-221/Domain: EGF homology <EGF>

```
Query Match          37.0%;   Score 102.5;  DB 2;   Length 636;
Best Local Similarity  31.2%;   Pred. No. 0.00038;
Matches   15;  Conservative 13;  Mismatches 19;  Indels  1;  Gaps  1;
```

```
Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
          |   | :|: |:|:| ||||  :|:  :|:    | :|:  |  | || ::
Db  178 HLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLCKCPNEFTGDRCQNYV 225
```

Search completed: November 7, 2002, 10:00:27
Job time : 19 secs

OM protein - protein search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: November 7, 2002, 09:59:04 ; Search time 30 Seconds
(without alignments)
174.016 Million cell updates/sec

Title: US-09-480-977-4
Perfect score: 277
Sequence: 1 HFKPCRDKDLAYCLNDGECF..........SHKHCRCKEGYQGVRCDQFL 47

Scoring table: BLOSUM62
    Gapop 10.0 , Gapext 0.5
Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
    Maximum Match 100%
    Listing first 45 summaries

Database :
    1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
    2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
    3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
    4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
    5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
    6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
    7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
    8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
    9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
    10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
    11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
    12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
    13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
    14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
    15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
    16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
    17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
    18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
    19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
    20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
    21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
    22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
    A:Geneseq-032802:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	277	100.0	47	20	AAW97622	Human neuregulin r
2	277	100.0	157	20	AAY05451	Human heregulin-11
3	277	100.0	360	20	AAW97621	Human heregulin-11
4	277	100.0	362	20	AAW97620	Mouse neuregulin r
5	277	100.0	696	20	AAW97619	Mouse neuregulin r
6	277	100.0	713	20	AAW97617	Mouse neuregulin r
7	277	100.0	720	20	AAY05452	Human neuregulin r
8	277	100.0	720	20	AAY05418	Human heregulin-11
9	116.5	42.1	52	17	AAW97618	Human neuregulin r
10	116.5	42.1	52	21	AAB12602	Human NDF EGF-like
11	116.5	42.1	52	21	AAY69983	NDF/heregulin prot

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

12	113.5	41.0	52	17	AAW05184	Neu differentiatio
13	113.5	41.0	63	15	AAR46918	EGFL2. AAR46918;
14	113.5	41.0	63	15	AAR55659	EGFL2. AAR55659;
15	113.5	41.0	63	16	AAR67250	Human epidermal 11
16	113.5	41.0	63	17	AAW09363	EGFL2. Not specif
17	113.5	41.0	63	17	AAR96076	Epidermal growth f
18	113.5	41.0	63	17	AAR87461	Epidermal growth f
19	113.5	41.0	66	21	AAR36702	EGF-like domain of
20	113.5	41.0	83	15	AAR46922	EGFL6. AAR46922;
21	113.5	41.0	83	15	AAR55663	EGFL6. AAR55663;
22	113.5	41.0	83	16	AAR67254	Human epidermal 11
23	113.5	41.0	83	17	AAW09367	EGFL6. Not specif
24	113.5	41.0	83	17	AAR96080	Epidermal growth f
25	113.5	41.0	83	17	AAR87465	Epidermal growth f
26	113.5	41.0	88	15	AAR46921	EGFL5. AAR46921;
27	113.5	41.0	88	15	AAR55662	EGFL5. AAR55662;
28	113.5	41.0	88	16	AAR67253	Human epidermal 11
29	113.5	41.0	88	17	AAW09366	EGFL5. Not specif
30	113.5	41.0	88	17	AAR96079	Epidermal growth f
31	113.5	41.0	88	17	AAR87464	Epidermal growth f
32	113.5	41.0	101	22	AAR67933	Human NRG1 protein
33	113.5	41.0	101	22	AAR67971	Human NRG1 protein
34	113.5	41.0	125	16	AAR68564	Human NDF-alpha1 c
35	113.5	41.0	125	16	AAR67928	Human NRG1 gene Ne
36	113.5	41.0	125	22	AAC67966	Human NRG1 gene Ne
37	113.5	41.0	263	13	AAR28537	GGF2BPP2.CDS prote
38	113.5	41.0	263	15	AAR46896	GGF2BPP2. Bos tau
39	113.5	41.0	263	15	AAR55689	GGF2BPP2. Bos tau
40	113.5	41.0	263	16	AAR67217	Putative bovine gl
41	113.5	41.0	280	17	AAW09360	Bovine neuregulin
42	113.5	41.0	280	15	AAR46915	GGF2BPP2. Bos tau
43	113.5	41.0	280	15	AAR55656	GGF2BPP2. Bos tau
44	113.5	41.0	280	16	AAR67244	Bovine glial cell
45	113.5	41.0	280	17	AAW09369	Human neuregulin G

ALIGNMENTS

RESULT 1
AAW97622
AAW97622 standard; Protein; 47 AA.

XX
ID   AAW97622 standard; Protein; 47 AA.
XX
AC   AAW97622;
XX
DT   10-MAY-1999 (first entry)
XX
DE   Human neuregulin related ligand NRG3 EGF-like domain.
XX
KW   Neuregulin related ligand; NRG3; hNRG3B1; ErbB4 receptor;
KW   signal transduction; nervous system disorder; neurodegeneration;
KW   neuropathy; therapy; diagnosis; epidermal growth factor; EGF;
KW   Immunoadhesin.
XX
OS   Homo sapiens.
XX
PN   WO9902681-A1.
XX
PD   21-JAN-1999.
XX
PF   30-JUN-1998; 98WO-US13411.
XX
PR   24-JUL-1997; 97US-0899437.
PR   09-JUL-1997; 97US-0052019.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Godowski PJ, Mark MR, Zhang D;
XX
DR   WPI; 1999-120882/10.
XX
PT   New isolated neuregulin related ligand-3 - used to develop products
PT   for treating nervous system disorders, e.g. stroke, ischaemia,

PT infection, malignancy, Alzheimer's disease or Down's syndrome
PS Claim 30; Page 64; 101pp; English.
XX
CC This is the epidermal growth factor (EGF)-like domain of human
CC neuregulin related ligand NRG3 (see also AAW97618), a novel member of
CC the EGF-like family of protein ligands that binds to the ErbB4
CC receptor and activates ErbB4 receptor tyrosine phosphorylation.
CC The EGF-1 like domain of NRG3 is distinct from the EGF-like domains
CC of NRG1 and NRG2. The invention provides human and murine
CC polypeptides (see also AAW97617) that have at least 75% homology to
CC the NRG3 EGF-like domain, as well as expression vectors, host cells
CC and methods for the recombinant production of novel NRG3s. The
CC NRG3 polypeptides and polynucleotides and can be used to enhance
CC the survival, proliferation or differentiation of cells having the
CC ErbB4 receptor in vivo and in vitro. They can be used to prevent
CC or treat damage to a nerve or damage to other NRG3-expressing or
CC NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC particular, they can be used to treat diseases which involve neural
CC cell growth such as demyelination, or damage or loss of glial cells
CC (e.g. multiple sclerosis). They can be used to treat patients whose
CC nervous system has been damaged by e.g. trauma, surgery, stroke,
CC ischaemia, infection, metabolic disease, nutritional deficiency,
CC malignancy, or toxic agents. NRG3 can also be used to treat
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC Gehrig's disease), Bell's palsy, conditions involving spinal
CC muscular atrophy or paralysis, neurodegenerative disorders such as
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC and Meniere's disease. They can also be used to treat neuropathies
CC associated with systemic disease including post-polio syndrome,
CC hereditary neuropathies including Charcot-Marie-Tooth disease,
CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC disease, metachromatic leukodystrophy, Fabry's disease and
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC smooth muscle, such as muscular dystrophy or diseases caused by
CC skeletal or smooth muscle wasting. The products can also be used
CC for detection, diagnosis, for the production of transgenic or
CC knockout animals or for drug screening. A claimed immunoadhesin
CC comprises the human NRG3 EGF-like domain fused to an immunoglobulin
CC sequence.
XX
SQ Sequence 47 AA;

Query Match 100.0%; Score 277; DB 20; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1 HFKPCRDKDLAYCLNDGECFVIEILTGSHKHCRCKEGYQGVRCDQFL 47
       |||||||||||||||||||||||||||||||||||||||||||||||
Db  1 HFKPCRDKDLAYCLNDGECFVIEILTGSHKHCRCKEGYQGVRCDQFL 47
```

RESULT 2
AAY05451
ID AAY05451 standard; Protein; 157 AA.
XX
AC AAY05451;
XX
DT 06-JUL-1999 (first entry)
XX
DE Human heregulin-like factor sequence.
XX
KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;
KW neural system disorder; cancer.
OS Homo sapiens.
XX
PN WO9857989-A1.
XX
PD 23-DEC-1998.
XX
PF 16-JUN-1998; 98WO-US12403.

PR 17-JUN-1997; 97US-0049942.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (GEOU) UNIV GEORGETOWN.
XX
PI Hijazi MM, King CR, Ruben SM, Young P;
XX
DR WPI; 1999-095327/08.
DR N-PSDB; AAX36423.
XX
PT New isolated heregulin-like factor - used to develop products for
PT the diagnosis and treatment of disorders involving regulation of
PT cell growth, particularly cancers
XX
PS Claim 17; Page 86-87; 118pp; English.
XX
CC This sequence is the human heregulin-like factor (HLF) of the
CC invention. The HLF is involved in the regulation of cell growth.
CC Detection of different levels of expression of the HLF gene can be used
CC for the diagnosis of disorders, e.g. in the neural system. In
CC particular, detection of different levels of HLF gene expression in cells
CC or body fluid of an individual can be used for diagnosing cancer. The
CC products can also be used in the treatment of disorders involving
CC abnormal levels of HLF activity.
XX
SQ Sequence 157 AA;

Query Match 100.0%; Score 277; DB 20; Length 157;
Best Local Similarity 100.0%; Pred. No. 7.6e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  31 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 77
```

RESULT 3
AAW97621
ID AAW97621 standard; Protein; 360 AA.
XX
AC AAW97621;
XX
DT 10-MAY-1999 (first entry)
XX
DE Human neuregulin related ligand NRG3 extracellular domain.
XX
KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW signal transduction; nervous system disorder; neurodegeneration;
KW neuropathy; therapy; diagnosis.
XX
PA (GETH) GENENTECH INC.
XX
PI Godowski PJ, Mark MR, Zhang D;
XX
DR WPI; 1999-120882/10.
XX
PN WO9902681-A1.
XX
PD 21-JAN-1999.
XX
PF 30-JUN-1998; 98WO-US13411.
XX
PR 24-JUL-1997; 97US-0899437.
PR 09-JUL-1997; 97US-0052019.
XX
OS Homo sapiens.

PT New isolated neuregulin related ligand-3 - used to develop products
PT for treating nervous system disorders, e.g. stroke, ischaemia,
PT infection, malignancy, Alzheimer's disease or Down's syndrome
PS Claim 5(a); Page 69-70; 101pp; English.

Left portion

```
XX
CC   This is the extracellular domain (ECD, aa1-360 of human neuregulin
CC   related ligand NRG3 (see also AAW97618), a novel member of the
CC   epidermal growth factor (EGF)-like family of protein ligands. NRG3
CC   binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,
CC   activates ErbB4 receptor tyrosine phosphorylation. The invention
CC   provides human and murine polypeptides (see also AAW97617) that have
CC   at least 75% homology to the NRG3 ECD, as well as expression vectors,
CC   host cells and methods for the recombinant production of novel
CC   NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
CC   enhance the survival, proliferation or differentiation of cells
CC   having the ErbB4 receptor in vivo and in vitro. They can be used to
CC   prevent or treat damage to a nerve or damage to other NRG3-expressing
CC   or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC   particular, they can be used to treat diseases which involve neural
CC   cell growth such as demyelination, or damage or loss of glial cells
CC   (e.g. multiple sclerosis). They can be used to treat patients whose
CC   nervous system has been damaged by e.g. trauma, surgery, stroke,
CC   ischaemia, infection, metabolic disease, nutritional deficiency,
CC   malignancy, or toxic agents. NRG3 can also be used to treat
CC   motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC   Gehrig's disease), Bell's palsy, conditions involving spinal
CC   muscular atrophy or paralysis, neurodegenerative disorders such as
CC   Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC   sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC   and Meniere's disease. They can also be used to treat neuropathies
CC   associated with systemic disease including Charcot-Marie-Tooth
CC   disease, metachromatic leukodystrophy, Fabry's disease and
CC   Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC   Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC   hereditary neuropathies including Charcot-Marie-Tooth disease,
CC   smooth muscle, such as muscular dystrophy or diseases caused by
CC   skeletal or smooth muscle wasting. The products can also be used
CC   knockout animals or for drug screening.
CC   for detection, diagnosis, for the production of transgenic or
XX
SQ   Sequence       360 AA;

Query Match             100.0%;  Score 277;  DB 20;  Length 360;
Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
Matches  47;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332
```

RESULT 4
```
ID   AAW97620 standard; Protein; 362 AA.
XX
AC   AAW97620;
XX
DT   10-MAY-1999 (first entry)
XX
DE   Mouse neuregulin related ligand NRG3 extracellular domain.
XX
OS   Mus sp.
XX
PN   WO9902681-A1.
XX
PD   21-JAN-1999.
XX
PF   30-JUN-1998;  98WO-US13411.
XX
PR   24-JUL-1997;  97US-0899437.
PR   09-JUL-1997;  97US-0052019.
XX
PA   (GETH ) GENENTECH INC.
```

Right portion

```
XX
PI   Godowski PJ, Mark MR, Zhang D;
XX
DR   WPI; 1999-120882/10.
XX
PT   New isolated neuregulin related ligand-3 - used to develop products
PT   for treating nervous system disorders, e.g. stroke, ischaemia,
PT   infection, malignancy, Alzheimer's disease or Down's syndrome
XX
PS   Claim 5(a); Page 62-63; 101pp; English.
```

```
CC   This is the extracellular domain (ECD, aa1-362) of murine neuregulin
CC   related ligand NRG3 (see also AAW97617), a novel member of the
CC   epidermal growth factor (EGF)-like family of protein ligands. NRG3
CC   binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,
CC   activates ErbB4 receptor tyrosine phosphorylation. The invention
CC   provides human and murine polypeptides (see also AAW97618) that have
CC   at least 75% homology to the NRG3 ECD, as well as expression vectors,
CC   host cells and methods for the recombinant production of novel
CC   NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
CC   enhance the survival, proliferation or differentiation of cells
CC   having the ErbB4 receptor in vivo and in vitro. They can be used to
CC   prevent or treat damage to a nerve or damage to other NRG3-expressing
CC   or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC   particular, they can be used to treat diseases which involve neural
CC   cell growth such as demyelination, or damage or loss of glial cells
CC   (e.g. multiple sclerosis). They can be used to treat patients whose
CC   nervous system has been damaged by e.g. trauma, surgery, stroke,
CC   ischaemia, infection, metabolic disease, nutritional deficiency,
CC   malignancy, or toxic agents. NRG3 can also be used to treat
CC   motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC   Gehrig's disease), Bell's palsy, conditions involving spinal
CC   muscular atrophy or paralysis, neurodegenerative disorders such as
CC   Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC   sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC   and Meniere's disease. They can also be used to treat neuropathies
CC   associated with systemic disease including Charcot-Marie-Tooth
CC   disease, metachromatic leukodystrophy, Fabry's disease and
CC   Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC   Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC   hereditary neuropathies including Charcot-Marie-Tooth disease,
CC   smooth muscle, such as muscular dystrophy or diseases caused by
CC   skeletal or smooth muscle wasting. The products can also be used
CC   knockout animals or for drug screening.
CC   for detection, diagnosis, for the production of transgenic or
XX
SQ   Sequence       362 AA;

Query Match             100.0%;  Score 277;  DB 20;  Length 362;
Best Local Similarity   100.0%;  Pred. No. 1.6e-20;
Matches  47;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  288 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334
```

RESULT 5
```
ID   AAW97619 standard; Protein; 696 AA.
XX
AC   AAW97619;
XX
DT   10-MAY-1999 (first entry)
XX
DE   Human neuregulin related ligand NRG3 (splice variant).
XX
OS   Homo sapiens.
XX
KW   Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW   signal transduction; nervous system disorder; neurodegeneration;
KW   neuropathy; therapy; diagnosis; splice variant.
```

Left column:

```
FH  Key          Location/Qualifiers
FT  Domain       1..360
                 /note= "extracellular domain, specifically claimed
                 in Claim 5(a)."
FT  Region       66..91
                 /note= "hydrophobic region"
FT  Region       101..284
                 /note= "mucin-like Ser/Thr-rich region, contains
                 sites for O-linked glycosylation"
FT  Domain       285..354
                 /note= "EGF-like domain"
FT  Domain       356..394
                 /note= "transmembrane domain"
XX
PN  WO9902681-A1.
XX
PD  21-JAN-1999.
XX
PF  30-JUN-1998.
XX
PR  24-JUL-1997;  97US-0899437;
PR  09-JUL-1997;  97US-0052019.
XX
PS  Example 1: Page 78-81; 101pp; English.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Godowski PJ, Mark MR, Zhang D;
XX
DR  WPI; 1999-120882/10.
DR  N-PSDB; AAX06989.
XX
PT  New isolated neuregulin related ligand-3 - used to develop products
PT  for treating nervous system disorders, e.g. stroke, ischaemia,
PT  Infection, malignancy, Alzheimer's disease or Down's syndrome
XX
CC  This is the amino acid sequence of splice variant hNGR3B2 of human
CC  neuregulin related ligand NRG3, a novel member of the epidermal
CC  growth factor (EGF)-like family of protein ligands that binds to
CC  the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, and
CC  which activates ErbB4 receptor tyrosine phosphorylation. The
CC  sequence was deduced from the nucleotide sequence of a cDNA clone
CC  (see AAX06989) from a foetal brain library. hNGR3B2 lacks amino
CC  acids 529-552 of hNGR3B1 (see AAW97618) but retains the EGF- like
CC  domain and is expected to exhibit biological activity. The invention
CC  provides human and murine NRG3 polypeptides (see AAW97617), expression
CC  vectors, host cells and methods for the recombinant production of
CC  NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
CC  enhance the survival, proliferation or differentiation of cells
CC  having the ErbB4 receptor in vivo and in vitro. They can be used to
CC  prevent or treat damage to a nerve or damage to other NRG3-expressing
CC  or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC  particular, they can be used to treat diseases which involve neural
CC  cell growth such as demyelination, or damage or loss of glial cells
CC  (e.g. multiple sclerosis). They can be used to treat patients whose
CC  nervous system has been damaged by e.g. trauma, surgery, stroke,
CC  ischaemia, infection, metabolic disease, nutritional deficiency,
CC  malignancy, or toxic agents. NRG3 can also be used to treat
CC  motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC  Gehrig's disease), Bell's palsy, conditions involving spinal
CC  muscular atrophy or paralysis, neurodegenerative disorders such as
CC  Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC  sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC  and Meniere's disease. They can be used to treat neuropathies
CC  associated with systemic disease including post-polio syndrome,
CC  hereditary neuropathies including Charcot-Marie-Tooth disease,
CC  Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC  disease, metachromatic leukodystrophy, Fabry's disease and
CC  Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC  smooth muscle, such as muscular dystrophy or diseases caused by
CC  skeletal or smooth muscle wasting. The products can also be used
CC  for detection, diagnosis, for the production of transgenic or
CC  knockout animals or for drug screening.
```

Right column:

```
XX
SQ  Sequence    696 AA;

    Query Match            100.0%;  Score 277;  DB 20;  Length 696;
    Best Local Similarity  100.0%;  Pred. No. 3e-20;
    Matches   47;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy  1    HFKPCDRDKDLAYCLNDGECFVIEFLTGSHKHCRCKEGYQGVRCDQFL  47
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  286  HFKPCDRDKDLAYCLNDGECFVIEFLTGSHKHCRCKEGYQGVRCDQFL  332

RESULT 6
AAW97617
ID  AAW97617 standard; Protein; 713 AA.
XX
AC  AAW97617;
XX
DT  10-MAY-1999  (first entry)
XX
DE  Mouse neuregulin related ligand NRG3.
XX
KW  Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;
KW  signal transduction; nervous system disorder; neurodegeneration;
KW  neuropathy; therapy; diagnosis.
XX
OS  Mus sp.
XX
FH  Key          Location/Qualifiers
FT  Domain       1..362
                 /note= "extracellular domain, specifically claimed
                 in Claim 5(a)."
FT  Region       66..91
                 /note= "hydrophobic region"
FT  Region       105..286
                 /note= "mucin-like Ser/Thr-rich region, contains
                 sites for O-linked glycosylation"
FT  Domain       287..334
                 /note= "EGF-like domain"
FT  Domain       363..385
                 /note= "transmembrane domain"
XX
PN  WO9902681-A1.
XX
PD  21-JAN-1999.
XX
PF  30-JUN-1998.
XX
PR  24-JUL-1997;  97US-0899437.
PR  09-JUL-1997;  97US-0052019.
XX
PS  Claim 5(b); Page 59-62; 101pp; English.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Godowski PJ, Mark MR, Zhang D;
XX
DR  WPI; 1999-120882/10.
DR  N-PSDB; AAX06987.
XX
PT  New isolated neuregulin related ligand-3 - used to develop products
PT  for treating nervous system disorders, e.g. stroke, ischaemia,
PT  Infection, malignancy, Alzheimer's disease or Down's syndrome
XX
CC  This is the amino acid sequence of murine neuregulin related ligand
CC  NRG3, a novel member of the epidermal growth factor (EGF)-like
CC  family of protein ligands that binds to the ErbB4 receptor, but not
CC  to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor
CC  tyrosine phosphorylation. The sequence was deduced from the
CC  nucleotide sequences of cDNA clones (see AAX06987) from a mouse brain
CC  library. The EGF-like domain of NRG3 is distinct from those of NRG1
CC  or NRG2, and NRG3 displays receptor binding characteristics that are
CC  distinct from those of other neuregulins. The invention provides
```

CC   human and murine NRG3 polypeptides (see also AAW97618), expression
CC   vectors, host cells and methods for the recombinant production of
CC   NRG3s. The NRG3 polypeptides and polynucleotides and can be used to
CC   enhance the survival, proliferation or differentiation of cells
CC   having the ErbB4 receptor in vivo and in vitro. They can be used to
CC   prevent or treat damage to a nerve or damage to other NRG3-expressing
CC   or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC   particular, they can be used to treat diseases which involve neural
CC   cell growth such as demyelination, or damage or loss of glial cells
CC   (e.g. multiple sclerosis). They can be used to treat patients whose
CC   nervous system has been damaged by e.g. trauma, surgery, stroke,
CC   ischaemia, infection, metabolic disease, nutritional deficiency,
CC   malignancy, or toxic agents. NRG3 can also be used to treat
CC   motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC   Gehrig's disease), Bell's palsy, conditions involving spinal
CC   muscular atrophy or paralysis, neurodegenerative disorders such as
CC   Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC   sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC   and Meniere's disease. They can also be used to treat neuropathies
CC   associated with systemic disease including post-polio syndrome,
CC   hereditary neuropathies including Charcot-Marie-Tooth disease,
CC   Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC   disease, metachromatic leukodystrophy, Fabry's disease and
CC   Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC   smooth muscle, such as muscular dystrophy or diseases caused by
CC   skeletal or smooth muscle wasting. The products can also be used
CC   for detection, diagnosis, for the production of transgenic or
CC   knockout animals or for drug screening.
XX
SQ   Sequence 713 AA;

Qy   1   HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   288 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334

Query Match   100.0%;   Score 277;   DB 20;   Length 713;
Best Local Similarity 100.0%;   Pred. No. 3.1e-20;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7
AAY05452
ID   AAY05452 standard; Protein; 720 AA.
XX
AC   AAY05452;
XX
DT   06-JUL-1999 (first entry)
XX
DE   Human heregulin-like factor sequence.
XX
KW   Human heregulin-like factor; HLF; cell growth regulator; diagnosis;
KW   neural system disorder; cancer.
XX
OS   Homo sapiens.
XX
PN   WO985789-A1.
XX
PD   23-DEC-1998.
XX
PR   16-JUN-1998; 98WO-US12403.
PR   17-JUN-1997; 97US-0049942.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (GEOU) UNIV GEORGETOWN.
XX
PI   Hjazi MM, King CR, Ruben SM, Young P;
XX
DR   WPI; 1999-095327/08.
XX
PT   New isolated heregulin-like factor - used to develop products for
PT   the diagnosis and treatment of disorders involving regulation of
PT   cell growth, particularly cancers

XX
PS   Disclosure; Page 97-99; 118pp; English.
XX
CC   This sequence is the human heregulin-like factor (HLF) of the
CC   invention. The HLF is involved in the regulation of cell growth.
CC   Detection of different levels of expression of the HLF gene can be
CC   for the diagnosis of disorders, e.g. in the neural system. In
CC   particular, detection of different levels of HLF gene expression in cells
CC   or body fluid of an individual can be used for diagnosing cancer. The
CC   products can also be used in the treatment of disorders involving
CC   abnormal levels of HLF activity.
XX
SQ   Sequence 720 AA;

Query Match   100.0%;   Score 277;   DB 20;   Length 720;
Best Local Similarity 100.0%;   Pred. No. 3.1e-20;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1   HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332

RESULT 8
AAW97618
ID   AAW97618 standard; Protein; 720 AA.
XX
AC   AAW97618;
XX
DT   10-MAY-1999 (first entry)
XX
DE   Human neuregulin related ligand NRG3.
XX
KW   Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;
KW   signal transduction; nervous system disorder; neurodegeneration;
KW   neuropathy; therapy; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO9902681-A1.
XX
PD   21-JAN-1999.
XX
PF   30-JUN-1998; 98WO-US13411.
XX
PR   24-JUL-1997; 97US-0899437.
PR   09-JUL-1997; 97US-0052019.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Godowski PJ, Mark MR, Zhang D;
XX
DR   WPI; 1999-120882/10.
DR   N-PSDB; AAX06908.
XX
PT   New isolated neuregulin related ligand-3 - used to develop products
PT   for treating nervous system disorders, e.g. stroke, ischaemia,
PT   infection, malignancy, Alzheimer's disease or Down's syndrome
XX

FH   Key   Location/Qualifiers
FH
FT   Domain   1..360
FT   /note= "extracellular domain,-specifically claimed
FT   in Claim 5(a)"
FT   Region   66..91
FT   /note= "hydrophobic region"
FT   Region   101..284
FT   /note= "mucin-like Ser/Thr-rich region, contains
FT   sites for O-linked glycosylation"
FT   Domain   285..354
FT   /note= "EGF-like domain"
FT   Domain   356..394
FT   /note= "transmembrane domain"
XX

```
PS   Claim 5(b); Page 66-69; 101pp; English.
XX
CC   This is the amino acid sequence of human neuregulin related ligand
CC   NRG3, a novel member of the epidermal growth factor (EGF)-like
CC   family of protein ligands that binds to the ErbB4 receptor, but not
CC   to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor
CC   tyrosine phosphorylation. The sequence was deduced from the
CC   nucleotide sequence of a cDNA clone (see AAX05988) from a foetal brain
CC   library. The EGF-like domain of NRG3 is distinct from those of NRG1
CC   or NRG2, and NRG3 displays receptor binding characteristics that are
CC   distinct from those of other neuregulins. An alternatively spliced
CC   form of human NRG3 is provided in AAW97619. The invention provides
CC   human and murine NRG3 polypeptides (see also AAW97617), expression
CC   vectors, host cells and methods for the recombinant production of
CC   NRG3s. The NRG3 polypeptides and polynucleotides can be used to
CC   enhance the survival, proliferation or differentiation of cells
CC   having the ErbB4 receptor in vivo and in vitro. They can be used to
CC   prevent or treat damage to a nerve or damage to other NRG3-expressing
CC   or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In
CC   particular, they can be used to treat diseases which involve neural
CC   cell growth such as demyelination, or damage or loss of glial cells
CC   (e.g. multiple sclerosis). They can be used to treat patients whose
CC   nervous system has been damaged by e.g. trauma, surgery, stroke,
CC   ischaemia, infection, metabolic disease, nutritional deficiency,
CC   malignancy, or toxic agents. NRG3 can also be used to treat
CC   motor neuron disorders such as amyotrophic lateral sclerosis (Lou
CC   Gehrig's disease), Bell's palsy, conditions involving spinal
CC   muscular atrophy or paralysis, neurodegenerative disorders such as
CC   Alzheimer's disease, Parkinson's disease, epilepsy, multiple
CC   sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,
CC   and Meniere's disease. They can also be used to treat neuropathies
CC   associated with systemic disease including post-polio syndrome,
CC   hereditary neuropathies including Charcot-Marie-Tooth disease,
CC   Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's
CC   disease, metachromatic leukodystrophy, Fabry's disease and
CC   Dejerine-Sottas syndrome, to treat disease of skeletal muscle of
CC   smooth muscle, such as muscular dystrophy or diseases caused by
CC   skeletal or smooth muscle wasting. The products can also be used
CC   for detection, diagnosis, for the production of transgenic or
CC   knockout animals or for drug screening.
XX
SQ   Sequence   720 AA;

Query Match             100.0%;  Score 277;  DB 20;  Length 720;
Best Local Similarity   100.0%;  Pred. No. 3.1e-20;
Matches    47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL   47
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   286  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL  332


RESULT 9
AAW05182
ID   AAW05182 standard; peptide; 52 AA.
XX
AC   AAW05182;
XX
DT   04-JUN-1997  (first entry)
XX
DE   Neu differentiation factor/heregulin-alpha/beta form EGF-like domain.
XX
KW   NDF; neu differentiation factor; heregulin; epidermal growth factor;
KW   EGF; colon epithelial cell proliferation; Schwann cell; nerve;
KW   damage; colitis; ulcer.
OS   Synthetic.
XX
PN   WO9631599-A1.
XX
PD   10-OCT-1996.
XX
PF   27-MAR-1996;  96WO-US04262.
XX
PR   06-APR-1995;  95US-0417640.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Cranahan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;
XX
PT   WPI; 1996-465022/46.
XX
PS   Claim 1; Page 27; 37pp; English.
XX
CC   The peptides AAW05182-W05185 are based on neu differentiation factor
CC   (NDF)/heregulin alpha and beta form EGF-like domains in various
CC   combinations. The peptides maintain the survival and proliferation of
CC   Schwann cells and cause proliferation, growth and differentiation of
CC   colon epithelial cells. Accordingly, they are useful to treat (in vitro
CC   or in vivo) a disease or disorder of the colon (e.g. colitis or an
CC   ulcer) or of the nervous system (e.g. nerve damage caused by trauma).
XX
SQ   Sequence   52 AA;

Query Match              42.1%;  Score 116.5;  DB 17;  Length 52;
Best Local Similarity    33.3%;  Pred. No. 4.7e-05;
Matches    16;  Conservative   15;  Mismatches   16;  Indels    1;  Gaps    1;

Qy     1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL   47
           | |:|| |||||: | : | ||||| ||||| |:|: | | || ::
Db     2  HLVRCAEEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCQNYV   49


RESULT 10
AAB12602
ID   AAB12602 standard; Peptide; 52 AA.
XX
AC   AAB12602;
XX
DT   09-NOV-2000  (first entry)
XX
DE   Human NDF EGF-like domain derived peptide SEQ ID NO:1.
XX
KW   Human; sensory epithelial cell; growth; stimulant; inner ear; EGF;
KW   epithelial growth factor; NDF; heregulin; monoclonal antibody;
KW   adult rat utricular epithelium.
OS   Homo sapiens.
XX
PN   US6080845-A.
XX
PD   27-JUN-2000.
XX
PF   28-JAN-1999;  99US-0238182.
XX
PR   05-AUG-1998;  98US-0129549.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Carnahan JF;
XX
DR   WPI; 2000-451229/39;
DR   N-PSDB; AAA60998.
XX
PT   Novel monoclonal antibody against adult rat utricular epithelium useful
PT   in study and research with such organs and tissue
XX
PS   Example 1; Fig 1; 12pp; English.
XX
CC   The present invention describes a monoclonal antibody (I) deposited with
CC   the American Type Culture Collection under accession number HB-12598.
CC   The antibodies are used in study and research with adult rat utricular
```

CC   organs and tissues. The present sequence is a peptide derived from
CC   the epithelial growth factor (EGF) like domains of NDF-alpha and
CC   NDF-beta (members of the NDF/heregulin protein family, which is used in
CC   the exemplification of the present invention. The peptide acts as a
CC   growth stimulant for sensory epithelial cells of the inner ear.
XX
SQ   Sequence   52 AA;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
      | :|: |||:: |||||::: || : | |:| | || :|
Db   2 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCQNYV 49

Query Match   42.1%;   Score 116.5;   DB 21;   Length 52;
Best Local Similarity   33.3%;   Pred. No. 4.7e-05;
Matches   16;   Conservative   15;   Mismatches   16;   Indels   1;   Gaps   1;

RESULT 11
AAY69983
ID   AAY69983   standard; peptide; 52 AA.
XX
AC   AAY69983;
XX
DT   20-APR-2000   (first entry)
XX
DE   NDF/heregulin protein family derivative peptide.
XX
KW   NDF/heregulin protein family; growth factor; proliferation stimulator;
KW   sensory epithelial cell; inner ear; vestibular disorder; hearing loss;
KW   balance disorder; therapy.
XX
OS   Synthetic.
XX
PN   US6017886-A.
XX
PD   Carnahan JF;
XX
PI   Carnahan JF;
XX
DR   WPI; 2000-136686/12.
XX
PT   Stimulating the proliferation of sensory epithelial cells of the inner
PT   ear with a peptide, useful for treating hearing loss -
XX
PS   Claim 1; Fig 1; 11pp; English.
XX
PR   05-AUG-1998;   98US-0129549.
XX
PF   23-FEB-1999;   99US-0255974.
XX
PD   25-JAN-2000.
XX
PN   US6017886-A.
XX
PA   (AMGE-) AMGEN INC.

ID   AAW05184   standard; peptide; 52 AA.
XX
AC   AAW05184;
XX
DT   04-JUN-1997   (first entry)
XX
DE   Neu differentiation factor/heregulin-alpha form EGF-like domain.
XX
KW   NDF; neu differentiation factor; heregulin; epidermal growth factor;
KW   EGF; colon epithelial cell proliferation; Schwann cell; nerve;
KW   damage; colitis; ulcer.
XX
OS   Synthetic.
XX
PN   WO9631599-A1.
XX
PD   10-OCT-1996.
XX
PF   27-MAR-1996;   96WO-US04262.
XX
PR   06-APR-1995;   95US-0417640.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Cranahan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;
XX
DR   WPI; 1996-465022/46.
XX
PT   Peptide(s) derived from neu differentiation factor/heregulin
PT   proteins - specifically from epidermal growth factor-like domain,
PT   stimulate proliferation of colon epithelial cells and Schwann cells
XX
PS   Disclosure; Page 24; 37pp; English.
XX
CC   The peptides AAW05182-W05185 are based on neu differentiation factor
CC   (NDF)/heregulin alpha and beta form EGF-like domains in various
CC   combinations. The peptides maintain the survival and proliferation of
CC   Schwann cells and cause proliferation, growth and differentiation of
CC   colon epithelial cells. Accordingly, they are useful to treat (in vitro
CC   or in vivo) a disease or disorder of the colon (e.g. colitis or an
CC   ulcer) or of the nervous system (e.g. nerve damage caused by trauma).
XX
SQ   Sequence   52 AA;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
      | :|: |||:: |||||::: |: : |: |:|: | || :
Db   2 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 47

Query Match   41.0%;   Score 113.5;   DB 17;   Length 52;
Best Local Similarity   34.8%;   Pred. No. 9.4e-05;
Matches   16;   Conservative   14;   Mismatches   15;   Indels   1;   Gaps   1;

RESULT 13
AAR46918
ID   AAR46918   standard; Protein; 63 AA.
XX
AC   AAR46918;
XX
DT   28-JUL-1994   (first entry)
XX
DE   EGFL2.
XX
KW   Glial growth factor; GGF; heregulin; mitogenesis;
KW   Schwann cell; tumour; central nervous system; erbB2 receptor;
KW   antiproliferative; epidermal growth factor; EGF.
XX
PN   WO9403644-A.
XX
PD   17-FEB-1994.
XX
PF   10-AUG-1993;   93WO-US07491.
XX

RESULT 12
AAW05184
Query Match   42.1%;   Score 116.5;   DB 21;   Length 52;
Best Local Similarity   33.3%;   Pred. No. 4.7e-05;
Matches   16;   Conservative   15;   Mismatches   16;   Indels   1;   Gaps   1;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
      | :|: |||:: |||||::: || : | |:| | || :
Db   2 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCQNYV 49

Sequence   52 AA;

This sequence represents an NDF/heregulin protein family derivative
peptide. The invention relates to a method for stimulating the
proliferation of sensory epithelial cells of the inner ear, comprising
contacting the cells with this sequence. The peptide acts as a growth
factor to stimulate proliferation of cells in the sensory epithelium. The
method is useful for the treatment of vestibular disorders (i.e. a
balance disorder) or for treating hearing loss.

SQ   Sequence   52 AA;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
      | :|: |||:: |||||::: || : | |:| | || :
Db   2 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCQNYV 49

```
PR   10-AUG-1992;   92US-0927337.
PR   25-SEP-1992;   92US-0951747.
PR   01-DEC-1992;   92US-0984085.
PR   29-JAN-1993;   93US-0011396.
XX
PA   (CAMB-) CAMBRIDGE NEUROSCIENCE INC.
XX
PI   Gwynne DI, Marchionni M, McBurney RN;
XX
DR   WPI; 1994-065731/08.
DR   N-PSDB; AAQ58324.
XX
PT   Glial growth factor DNA encoding numerous polypeptide factors
PT   used for inhibiting cell proliferation - for treating carcinoma
PT   and nervous disorders
XX
PS   Disclosure; Fig 40; 178pp; English.
XX
CC   The GGF coding segments include regions with EGF-like homology.
CC   These EGF-like domains can be required for the activation of
CC   mitogenesis in the binding reaction between GGF ligands contg.
CC   such domains and the erbB2 receptor. Pref. antiproliferative
CC   factors are those which lack these EGF-like domains.
XX
SQ   Sequence   63 AA;

Qy   1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
        | :|: |||||:|||:|:|: |:|: |:| |||:
Db   2  HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 47

Query Match             41.0%;  Score 113.5;  DB 15;  Length 63;
Best Local Similarity   34.8%;  Pred. No. 0.00011;
Matches  16;  Conservative  14;  Mismatches  15;  Indels  1;  Gaps  1;

RESULT 14
AAR55659
ID   AAR55659 standard; Protein; 63 AA.
XX
AC   AAR55659;
XX
DT   28-JUL-1994 (first entry)
XX
DE   EGFL2.
XX
KW   Glial growth factor; GGF; heregulin; mitogenesis;
KW   Schwann cell; tumour; central nervous system;
KW   epidermal growth factor; EGF.
XX
PN   WO9400140-A.
XX
PD   06-JAN-1994.
XX
PF   29-JUN-1993;   93WO-US06228.
XX
PR   30-JUN-1992;   92US-0907138.
PR   03-SEP-1992;   92US-0940389.
PR   23-OCT-1992;   92US-0965173.
PR   24-MAR-1993;   93US-0036555.
XX
PA   (CAMB-) CAMBRIDGE NEUROSCIENCE;
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen MS, Goodearl A, Hiles I, Marchioni M, Minghetti L;
PI   Stroobant P, Waterfield M;
XX
DR   WPI; 1994-025882/03.
DR   N-PSDB; AAQ62843.
XX
PT   Glial mitogenic polypeptide factors - useful for stimulating
PT   glial cell mitogenesis and treating glial cell tumours
XX
PS   Claim 53; Fig 39; 178pp; English.
```

```
XX
CC   EGFL1, EGFL2, EGFL3, EGFL4, EGFL5 and EGFL6 are used
CC   for the stimulation of glial cell mitogenesis in vivo
CC   and in vitro.
XX
SQ   Sequence   63 AA;

Qy   1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
        | :|: |||||:|||:|:|: |:|: |:| |||:
Db   2  HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 47

Query Match             41.0%;  Score 113.5;  DB 15;  Length 63;
Best Local Similarity   34.8%;  Pred. No. 0.00011;
Matches  16;  Conservative  14;  Mismatches  15;  Indels  1;  Gaps  1;

RESULT 15
AAR67250
ID   AAR67250 standard; Protein; 63 AA.
XX
AC   AAR67250;
XX
DT   15-AUG-1995 (first entry)
XX
DE   EGFL2.
XX
OS   Homo sapiens.
XX
DE   Human epidermal like growth factor 2 (EGFL2).
XX
KW   Epidermal like growth factor 2; mammalian muscle cell treatment;
KW   skeletal; cardiac; smooth; acetylcholine receptor deficiency;
KW   EGFL2.
XX
PN   WO9426298-A.
XX
PD   24-NOV-1994.
XX
PF   06-MAY-1994;   94WO-US05083.
XX
PR   06-MAY-1993;   93US-0059022.
PR   08-MAR-1994;   94US-0209204.
XX
PA   (CAMB-) CAMBRIDGE NEUROSCIENCE.
XX
PI   Gwynne DI, Marchionni M, Sklar R;
XX
DR   WPI; 1995-006353/01.
DR   N-PSDB; AAQ74915.
XX
PT   Treating mammalian muscle diseases and disorders - by admin. of
PT   GGF2 and other specified polypeptide(s) which bind the p185erbB2
PT   receptor.
XX
PS   Claim 34; Pages 148-149; 241pp; English.
XX
CC   AAQ74915 encodes AAR67250 human epidermal like growth factor 2 (EGFL2).
CC   The glial cell mitogenic activity of EGFL2 can be used to treat a
CC   variety of mammalian skeletal, cardiac and smooth muscle diseases,
CC   including acetylcholine receptor deficiency.
XX
SQ   Sequence   63 AA;

Qy   1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQ 45
        | :|: |||||:|||:|:|: |:|: |:| |||:
Db   2  HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCTE 47

Query Match             41.0%;  Score 113.5;  DB 16;  Length 63;
Best Local Similarity   34.8%;  Pred. No. 0.00011;
Matches  16;  Conservative  14;  Mismatches  15;  Indels  1;  Gaps  1;

Search completed: November  7, 2002, 09:59:44
Job time : 32 secs
```

```
                    GenCore version 5.1.3
            Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          November  7, 2002, 09:59:49 ; Search time 13 seconds
                                                (without alignments)
                                             88.308 Million cell updates/sec

Title:          US-09-480-977-4
Perfect score:  277
Sequence:       1 HFKPCRDKDLAYCLNDGECF.........SHKHCRCKEGYQGVRCDQFL 47

Scoring table:  BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:       231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters:    231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	277	100.0	47	3	US-08-899-437-4	Sequence 4, Appl
2	277	100.0	47	3	US-08-899-437-8	Sequence 8, Appl
3	277	100.0	47	4	US-09-126-121-4	Sequence 4, Appl
4	277	100.0	47	4	US-09-126-121-8	Sequence 8, Appl
5	277	100.0	360	3	US-08-899-437-7	Sequence 7, Appl
6	277	100.0	360	4	US-09-126-121-7	Sequence 7, Appl
7	277	100.0	362	3	US-08-899-437-3	Sequence 3, Appl
8	277	100.0	362	4	US-09-126-121-3	Sequence 3, Appl
9	277	100.0	696	3	US-08-899-437-23	Sequence 23, Appl
10	277	100.0	696	4	US-09-126-121-23	Sequence 23, Appl
11	277	100.0	713	3	US-08-899-437-2	Sequence 2, Appl
12	277	100.0	713	4	US-09-126-121-2	Sequence 2, Appl
13	277	100.0	720	3	US-08-899-437-6	Sequence 6, Appl
14	277	100.0	720	4	US-09-126-121-6	Sequence 6, Appl
15	116.5	42.1	52	1	US-08-417-640A-1	Sequence 1, Appl
16	116.5	42.1	52	3	US-08-760-815-1	Sequence 1, Appl
17	116.5	42.1	52	2	US-08-761-038-1	Sequence 1, Appl
18	116.5	42.1	52	3	US-09-238-182-1	Sequence 1, Appl
19	113.5	41.0	49	3	US-08-899-437-14	Sequence 14, Appl
20	113.5	41.0	49	4	US-09-126-121-14	Sequence 14, Appl
21	113.5	41.0	50	4	US-08-753-007A-12	Sequence 12, Appl
22	113.5	41.0	50	3	US-09-398-496-12	Sequence 12, Appl
23	113.5	41.0	52	1	US-08-417-640A-3	Sequence 3, Appl
24	113.5	41.0	52	3	US-08-760-815-3	Sequence 3, Appl
25	113.5	41.0	52	2	US-08-761-038-3	Sequence 3, Appl
26	113.5	41.0	54	1	US-08-179-481-111	Sequence 111, App
27	113.5	41.0	63	3	US-08-341-018-62	Sequence 62, Appl
28	113.5	41.0	63	4	US-08-470-335-221	Sequence 221, App
29	113.5	41.0	63	4	US-08-470-339-221	Sequence 221, App
30	113.5	41.0	63	1	US-07-847-743B-10	Sequence 10, Appl
31	113.5	41.0	66	3	US-08-456-201-10	Sequence 10, Appl
32	113.5	41.0	66	2	US-08-456-241-10	Sequence 10, Appl
33	113.5	41.0	66	4	US-09-020-880-2	Sequence 2, Appl1
34	113.5	41.0	66	5	PCT-US92-04295A-10	Sequence 10, Appl
35	113.5	41.0	66	1	US-08-456-201-14	Sequence 14, Appl
36	113.5	41.0	83	4	US-07-847-743B-14	Sequence 14, App
37	113.5	41.0	83	4	US-08-341-018-70	Sequence 70, Appl
38	113.5	41.0	83	3	US-08-470-335-225	Sequence 225, App
39	113.5	41.0	88	1	US-08-470-339-225	Sequence 225, App
40	113.5	41.0	88	4	US-08-341-018-68	Sequence 68, Appl
41	113.5	41.0	95	4	US-08-470-335-224	Sequence 224, App
42	113.5	41.0	95	1	US-07-847-743B-14	Sequence 14, Appl
43	113.5	41.0	95	2	US-08-330-161-12	Sequence 12, Appl
44	113.5	41.0	95	2	US-08-456-241-14	Sequence 14, Appl
45	113.5	41.0	95	2	US-08-440-401-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-08-899-437-4
; Sequence 4, Application US/08899437
; Patent No. 6121415
; GENERAL INFORMATION:
;  APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;  TITLE OF INVENTION:  ErbB Receptor-Specific Neuregulin Related
;  TITLE OF INVENTION:  Ligands and Uses Therefor
;  NUMBER OF SEQUENCES:  23
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Genentech, Inc.
;   STREET:  1 DNA Way
;   CITY:  South San Francisco
;   STATE:  California
;   COUNTRY:  USA
;   ZIP:  94080
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  WinPatin (Genentech)
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/899,437
;   FILING DATE:  24-Jul-1997
;   CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;   NAME:  Conley, Deirdre L.
;   REGISTRATION NUMBER:  36,487
;   REFERENCE/DOCKET NUMBER:  P1084R1
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  650/225-2066
;   TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
;    LENGTH:  47 amino acids
;    TYPE:  Amino Acid
;    TOPOLOGY:  Linear
;  FEATURE:
;    NAME/KEY:  NRG3 EGF-like domain/amino acid seq.
;    LOCATION:  1-47
;    IDENTIFICATION METHOD:
;    OTHER INFORMATION:
US-08-899-437-4

Query Match           100.0%;  Score 277;  DB 3;  Length 47;
Best Local Similarity 100.0%;  Pred. No. 1.3e-26;
  Matches   47;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

QY   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
```

Db   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47

RESULT 2
US-08-899-437-8
Sequence 8, Application US/0889437
Patent No. 6121415

GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: NRG3 EGF-like domain/amino acid seq.
LOCATION: 1-47
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-8

Query Match   100.0%; Score 277; DB 3; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
        |||||||||||||||||||||||||||||||||||||||||||||||
Db   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47

RESULT 3
US-09-126-121-4
Sequence 4, Application US/09126121
Patent No. 6252051

GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:

Db   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: NRG3 EGF-like domain/amino acid seq.
LOCATION: 1-47
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-4

Query Match   100.0%; Score 277; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
        |||||||||||||||||||||||||||||||||||||||||||||||
Db   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47

RESULT 4
US-09-126-121-8
Sequence 8, Application US/09126121
Patent No. 6252051

GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear

```
; FEATURE:
;   NAME/KEY: NRG3 EGF-like domain/amino acid seq.
;   LOCATION: 1-47
;   IDENTIFICATION METHOD:
;   OTHER INFORMATION:
US-09-126-121-8

Qy   Query Match         100.0%; Score 277; DB 4; Length 47;
Db   Best Local Similarity 100.0%; Pred. No. 1.3e-26;
     Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
```

RESULT 5
US-08-899-437-7
```
; Sequence 7, Application US/0899437
; Patent No. 6121415
; GENERAL INFORMATION:
;   APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;   TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
;   TITLE OF INVENTION: Ligands and Uses Therefor
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 1 DNA Way
;     CITY: South San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/899,437
;     FILING DATE: 24-Jul-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Conley, Deirdre L.
;     REGISTRATION NUMBER: 36,487
;   REFERENCE/DOCKET NUMBER: P1084R1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 650/225-2066
;     TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 360 amino acids
;     TYPE: Amino Acid
;     TOPOLOGY: Linear
;   FEATURE:
;     NAME/KEY: hNRG3 extracellular domain/Amino AcidSeq
;     LOCATION: 1-360
;     IDENTIFICATION METHOD:
;     OTHER INFORMATION:
US-08-899-437-7

Qy   Query Match         100.0%; Score 277; DB 3; Length 360;
Db   Best Local Similarity 100.0%; Pred. No. 1e-25;
     Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332
```

RESULT 6
US-09-126-121-7
```
; Sequence 7, Application US/09126121
; Patent No. 6252051
; GENERAL INFORMATION:
;   APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;   TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
;   TITLE OF INVENTION: Ligands and Uses Therefor
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 1 DNA Way
;     CITY: South San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/126,121
;     FILING DATE: 30-Jul-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Conley, Deirdre L.
;     REGISTRATION NUMBER: 36,487
;   REFERENCE/DOCKET NUMBER: P1084R1D1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 650/225-2066
;     TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 360 amino acids
;     TYPE: Amino Acid
;     TOPOLOGY: Linear
;   FEATURE:
;     NAME/KEY: hNRG3 extracellular domain/Amino AcidSeq
;     LOCATION: 1-360
;     IDENTIFICATION METHOD:
;     OTHER INFORMATION:
US-09-126-121-7

Qy   Query Match         100.0%; Score 277; DB 4; Length 360;
Db   Best Local Similarity 100.0%; Pred. No. 1e-25;
     Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy    1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332
```

RESULT 7
US-08-899-437-3
```
; Sequence 3, Application US/0899437
; Patent No. 6121415
; GENERAL INFORMATION:
;   APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;   TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
;   TITLE OF INVENTION: Ligands and Uses Therefor
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 1 DNA Way
;     CITY: South San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/899,437
;     FILING DATE: 24-Jul-1997
```

```
·  CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
·    NAME: Conley, Deirdre L.
·    REGISTRATION NUMBER: 36,487
·    REFERENCE/DOCKET NUMBER: P1084R1
·    TELECOMMUNICATION INFORMATION:
·      TELEPHONE: 650/225-2066
·      TELEFAX: 650/952-9881
US-08-899-437-3
·  INFORMATION FOR SEQ ID NO: 3:
·    SEQUENCE CHARACTERISTICS:
·      LENGTH: 362 amino acids
·      TYPE: Amino Acid
·      TOPOLOGY: Linear
·    FEATURE:
·      NAME/KEY: mNRG3 extracellular domainAmino acid seq
·      LOCATION: 1-362
·      IDENTIFICATION METHOD:
·      OTHER INFORMATION:
US-08-899-437-3
```

```
Qy    1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  288  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334
```

```
Query Match          100.0%;   Score 277; DB 3;  Length 362;
Best Local Similarity 100.0%;  Pred. No. 1e-25;
Matches    47;  Conservative 0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
RESULT 8
US-09-126-121-3
;  Sequence 3, Application US/09126121
;  Patent No. 6252051
;  GENERAL INFORMATION:
;    APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;    TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
;    TITLE OF INVENTION: Ligands and Uses Therefor
;    NUMBER OF SEQUENCES: 23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Genentech, Inc.
;      STREET: 1 DNA Way
;      CITY: South San Francisco
;      STATE: California
;      COUNTRY: USA
;      ZIP: 94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/126,121
;      FILING DATE: 30-Jul-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Conley, Deirdre L.
;      REGISTRATION NUMBER: 36,487
;      REFERENCE/DOCKET NUMBER: P1084R1D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 650/225-2066
;      TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 362 amino acids
;      TYPE: Amino Acid
;      TOPOLOGY: Linear
;    FEATURE:
;      NAME/KEY: mNRG3 extracellular domainAmino acid seq
;      LOCATION: 1-362
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-126-121-3
```

```
Query Match          100.0%;   Score 277; DB 4;  Length 362;
Best Local Similarity 100.0%;  Pred. No. 1e-25;
Matches    47;  Conservative 0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy    1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  288  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334
```

```
RESULT 9
US-08-899-437-23
;  Sequence 23, Application US/0889437
;  Patent No. 6121415
;  GENERAL INFORMATION:
;    APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;    TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
;    TITLE OF INVENTION: Ligands and Uses Therefor
;    NUMBER OF SEQUENCES: 23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Genentech, Inc.
;      STREET: 1 DNA Way
;      CITY: South San Francisco
;      STATE: California
;      COUNTRY: USA
;      ZIP: 94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/899,437
;      FILING DATE: 24-Jul-1997
;      CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Conley, Deirdre L.
;      REGISTRATION NUMBER: 36,487
;      REFERENCE/DOCKET NUMBER: P1084R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 650/225-2066
;      TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 696 amino acids
;      TYPE: Amino Acid
;      TOPOLOGY: Linear
;    FEATURE:
;      NAME/KEY: Human NRG3B2
;      LOCATION: 1-696
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-899-437-23
```

```
Query Match          100.0%;   Score 277; DB 3;  Length 696;
Best Local Similarity 100.0%;  Pred. No. 2e-25;
Matches    47;  Conservative 0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy    1  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  286  HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332
```

```
RESULT 10
US-09-126-121-23
;  Sequence 23, Application US/09126121
;  Patent No. 6252051
;  GENERAL INFORMATION:
;    APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
;    TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
;    TITLE OF INVENTION: Ligands and Uses Therefor
;    NUMBER OF SEQUENCES: 23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Genentech, Inc.
```

STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
OTHER INFORMATION:
IDENTIFICATION METHOD:
US-09-126-121-23

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: Human NRG3B2
LOCATION: 1-696
IDENTIFICATION METHOD:
OTHER INFORMATION:

Qy
Db

Query Match   100.0%;   Score 277; DB 4; Length 696;
Best Local Similarity 100.0%;   Pred. No. 2e-25;
Matches   47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
  1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
    |||||||||||||||||||||||||||||||||||||||||||||||
286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332
```

RESULT 11
US-08-899-437-2
Sequence 2, Application US/0899437
Patent No. 6121415
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.
LOCATION: 1-713
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-2

Query Match   100.0%;   Score 277; DB 3; Length 713;
Best Local Similarity 100.0%;   Pred. No. 2.1e-25;
Matches   47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
Db

```
  1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
    |||||||||||||||||||||||||||||||||||||||||||||||
288 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334
```

RESULT 12
US-09-126-121-2
Sequence 2, Application US/09126121
Patent No. 6252051
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/952-9881
TELEFAX: 650/225-2066
OTHER INFORMATION:
IDENTIFICATION METHOD:
US-09-126-121-2

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.
LOCATION: 1-713
IDENTIFICATION METHOD:
OTHER INFORMATION:

Query Match   100.0%;   Score 277; DB 4; Length 713;
Best Local Similarity 100.0%;   Pred. No. 2.1e-25;
Matches   47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
Db

```
  1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
    |||||||||||||||||||||||||||||||||||||||||||||||
288 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 334
```

RESULT 13
US-08-899-437-6
Sequence 6, Application US/0899437
Patent No. 6121415
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
TITLE OF INVENTION: Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: hNRG3B1 amino acid sequence
LOCATION: 1-720
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-6

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332

Query Match   100.0%; Score 277; DB 3; Length 720;
Best Local Similarity   100.0%; Pred. No. 2.1e-25;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
US-09-126-121-6
Sequence 6, Application US/09126121
Patent No. 6252051
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
TITLE OF INVENTION: Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: hNRG3B1 amino acid sequence
LOCATION: 1-720
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-6

Qy   1 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 47
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 286 HFKPCRDKDLAYCLNDGECFVIETLTGSHKHCRCKEGYQGVRCDQFL 332

Query Match   100.0%; Score 277; DB 4; Length 720;
Best Local Similarity   100.0%; Pred. No. 2.1e-25;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
US-08-417-640A-1
Sequence 1, Application US/0841764OA
Patent No. 5670342
GENERAL INFORMATION:
APPLICANT: Carnahan, Josette F.
APPLICANT: Hara, Shinichi
APPLICANT: Lu, Hsieng S.
APPLICANT: Mayer, John P.
APPLICANT: Yoshinaga, Steven K.
TITLE OF INVENTION: NDF Peptides
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,640A
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REFERENCE/DOCKET NUMBER: A-310
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-417-640A-1

Query Match   42.1%; Score 116.5; DB 1; Length 52;
Best Local Similarity   33.3%; Pred. No. 1.9e-07;
Matches   16; Conservative   15; Mismatches   16; Indels   1; Gaps   1;

Qy   1 HPKPCRDKDLAYCLNDGECFVIETLTGSHKH-CRCKEGYQGVRCDQFL 47
           | : | :   : | ||| :: : |:|: | : | : | :| :: ::
Db   2 HLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLCKCQPGFTGARCQNYV 49

Search completed: November 7, 2002, 10:03:19
Job time : 14 secs